

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2002, 07:45:47 ; Search time 29.86 Seconds
(without alignments)
26.039 Million cell updates/sec

Title: US-09-641-801-1

Perfect score: 42

Sequence: 1 MPPPLP 7

Scoring table:

BLOSUM62

Gapop 10.0, Capext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : A.Genesec.032802.*

1:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1980.DAT.*
2:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1981.DAT.*
3:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1982.DAT.*
4:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1983.DAT.*
5:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1984.DAT.*
6:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1985.DAT.*
7:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1986.DAT.*
8:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1987.DAT.*
9:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1988.DAT.*
10:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1989.DAT.*
11:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1990.DAT.*
12:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1991.DAT.*
13:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1992.DAT.*
14:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1993.DAT.*
15:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1994.DAT.*
16:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1995.DAT.*
17:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1996.DAT.*
18:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1997.DAT.*
19:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1998.DAT.*
20:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1999.DAT.*
21:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA2000.DAT.*
22:	/SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	7	22	AA872246
2	42	100.0	7	22	AA872500
3	42	100.0	7	22	AA872532
4	42	100.0	7	22	AA859324
5	42	100.0	8	22	AA859354
6	42	100.0	316	22	AA867510
7	39	92.9	95	22	AA045697
8	39	92.9	132	22	ABB31347
9	39	92.9	132	22	ABB36552
10	39	92.9	132	22	ABB21889
11	39	92.9	132	22	AA857317

12	39	92.9	132	22	AA869716
13	39	92.9	132	22	AA817529
14	39	92.9	132	22	AA830051
15	39	92.9	132	22	AA805201
16	37	88.1	73	21	AAE10123
17	37	88.1	73	21	AAE11170
18	37	88.1	95	22	ABG21168
19	37	88.1	99	21	AA823978
20	37	88.1	278	16	AA879095
21	37	88.1	278	20	AA898069
22	37	88.1	278	20	AA895040
23	37	88.1	302	21	AA850933
24	37	88.1	332	22	AAE10122
25	37	88.1	393	22	AA883413
26	37	88.1	440	22	ABG60862
27	37	88.1	450	22	ABG09852
28	37	88.1	580	22	AA860493
29	37	88.1	586	22	AAE10121
30	37	88.1	667	19	AA848760
31	37	88.1	681	22	ABG07912
32	37	88.1	745	22	ABG58741
33	37	88.1	819	22	ABG67658
34	37	88.1	874	18	AA807600
35	37	88.1	874	20	AA894075
36	37	88.1	906	22	AA892816
37	37	88.1	908	22	AA808834
38	37	88.1	961	22	ABG59788
39	37	88.1	1003	22	AA879129
40	37	88.1	1217	22	AA833177
41	37	88.1	1244	22	AA867393
42	37	88.1	1343	22	AA800947
43	37	88.1	1359	21	AA841785
44	37	88.1	1475	22	AA871451
45	37	88.1	1554	22	AB811890
46	37	88.1	1629	22	ABG63122
47	37	88.1	1706	22	AA879270
48	37	88.1	3398	22	AAE03657
49	36	85.7	159	22	AAE03049
50	36	85.7	171	21	AA834163
51	36	85.7	239	22	ABG05698
52	36	85.7	282	22	ABG22406
53	36	85.7	282	22	AB831861
54	36	85.7	282	22	AB834678
55	36	85.7	282	22	AB837097
56	36	85.7	282	22	AB820088
57	36	85.7	282	22	AB822406
58	36	85.7	282	22	AA855473
59	36	85.7	282	22	AA857819
60	36	85.7	282	22	AA867858
61	36	85.7	282	22	AA870233
62	36	85.7	282	22	AA818059
63	36	85.7	282	22	AA828182
64	36	85.7	282	22	AA830571
65	36	85.7	282	22	AA803413
66	36	85.7	282	22	AA805699
67	36	85.7	318	22	AB864755
68	36	85.7	349	15	AA848750
69	36	85.7	349	17	AA802722
70	36	85.7	505	19	AA853808
71	36	85.7	505	22	AA852320
72	36	85.7	505	22	AA867360
73	36	85.7	655	22	ABG10987
74	36	85.7	818	22	AB870219
75	36	85.7	2234	22	ABG00985
76	35	83.3	57	21	AA818783
77	35	83.3	76	22	AB852686
78	35	83.3	91	21	AA842147
79	35	83.3	140	22	AA805152
80	35	83.3	183	22	AB866749
81	35	83.3	192	22	AB858711
82	35	83.3	202	22	ABG27240
83	35	83.3	233	21	AA854186
84	35	83.3	285	21	AA821494

Human bone marrow
Peptide #3963 enco
Peptide #4088 enco
Peptide #3883 enco
Human ion channel-
Human ORFX ORF934
Novel human diago
Atrialopis thalia
Rat Fas ligand enc
Rat Fas ligand (Fa
Human fetal brain
Human ion channel-
Human ion channel
Drosophila melanog
Novel human diago
Human cell cycle a
Human ion channel-
BOP1 protein. Mus
Novel human diago
Drosophila melanog
Drosophila melanog
Human G-protein re
Human G-protein co
Human protein sequ
Human bone marrow
Human bone marrow
Human ORFX ORF1549
Drosophila melanog
Human protocadheri
Drosophila melanog
Human protein SRO
Human extracellular
Human proctopile reg
Zee meys protein f
Novel human diago
Peptide #2154 enco
Peptide #4512 enco
Peptide #2184 enco
Peptide #4603 enco
Protein #2087 enco
Human brain expres
Human brain expres
Human bone marrow
Human bone marrow
Peptide #4493 enco
Peptide #2219 enco
Peptide #2608 enco
Peptide #4381 enco
Drosophila melanog
G-protein coupled
N-Wiskott-Aldrich
Bovine N-WASP prot
Novel acid sequenc
Novel human diago
Drosophila melanog
Novel human diago
Zee meys protein f
Zee meys protein f
Zee meys protein f
Human polypeptide
Drosophila melanog
Drosophila melanog
Novel human diago
Human pancreatic c
Atrialopis thalia

85	35	83.3	302	21	AA621493	Arabidopsis thalia
86	35	83.3	312	21	AA621492	Arabidopsis thalia
87	35	83.3	354	22	AA695571	Human protein sequ
88	35	83.3	412	22	AA692491	Human protein sequ
89	35	83.3	447	22	AA647151	Human 5-Ht3-C. Ho
90	35	83.3	531	21	AA656866	Human prostate can
91	35	83.3	555	20	AAV13455	Amino acid sequenc
92	35	83.3	555	20	AAV101085	Mammalian disabled
93	35	83.3	623	17	AA686875	Cancer suppressor
94	35	83.3	713	22	AA685905	Drosophila melanog
95	35	83.3	916	22	AA682553	Drosophila melanog
96	35	83.3	1485	22	AA640755	Human OREF519
97	35	83.3	1485	22	AA603458	Novel human diagno
98	35	83.3	1957	22	AA680145	Human protein sequ
99	35	83.3	2940	22	AA679161	Human protein sequ
100	34	81.0	15	21	AAV85491	Human ced-6 (hced-
101	34	81.0	16	21	AAV85494	Human ced-6 (hced-
102	34	81.0	28	21	AAV81249	Human Cdp (Cdc42hs
103	34	81.0	32	21	AAV81245	Human Cdp (Cdc42hs
104	34	81.0	65	21	AA626585	Arabidopsis thalia
105	34	81.0	67	22	AA684014	Human immune/hema
106	34	81.0	72	22	AAV6589	Human reproductive
107	34	81.0	76	20	AAV27251	Human hICED-6 prol
108	34	81.0	81	21	AA626584	Arabidopsis thalia
109	34	81.0	90	21	AA654723	Arabidopsis thalia
110	34	81.0	96	21	AA634178	Zea mays protein f
111	34	81.0	101	21	AA634177	Zea mays protein f
112	34	81.0	103	22	ABG18824	Novel human diagno
113	34	81.0	107	22	AAV59905	Human normal pancr
114	34	81.0	107	22	AA605620	Human polypeptide
115	34	81.0	111	22	ABG26028	Novel human diagno
116	34	81.0	112	22	AAV54568	Propionibacterium
117	34	81.0	114	22	AA626583	Arabidopsis thalia
118	34	81.0	117	22	AAV65524	Human reproductive
119	34	81.0	120	21	AA634176	Zea mays protein f
120	34	81.0	121	20	AAV66165	Human bladder tumo
121	34	81.0	121	21	AAV56893	Human prostate can
122	34	81.0	122	22	AAV56866	C. elegans UNC-5 p
123	34	81.0	123	22	AAV83911	Human immune/hema
124	34	81.0	127	22	AA603120	Human polypeptide
125	34	81.0	130	20	AAV36264	Human secreted pro
126	34	81.0	144	21	AAV99917	Peptide encoded by
127	34	81.0	144	21	AAV99918	Peptide encoded by
128	34	81.0	146	22	AAV92594	Human protein sequ
129	34	81.0	148	21	AA608181	Arabidopsis thalia
130	34	81.0	148	21	AAV52235	Arabidopsis thalia
131	34	81.0	152	21	AAV54970	Arabidopsis thalia
132	34	81.0	153	21	AAV27107	Zea mays protein f
133	34	81.0	155	21	AAV42721	Human OREF2485
134	34	81.0	164	21	AAV68180	Arabidopsis thalia
135	34	81.0	164	21	AAV52234	Arabidopsis thalia
136	34	81.0	174	22	AAV03637	Human polypeptide
137	34	81.0	178	22	AAV47551	Protonibacterium
138	34	81.0	194	19	AAV54363	Steroid membrane b
139	34	81.0	194	19	AAV39900	Pig plasma membran
140	34	81.0	195	19	AAV80396	A secreted protein
141	34	81.0	195	20	AAV08469	Human hSNMP1 prote
142	34	81.0	195	20	AAV88500	Human stomach carc
143	34	81.0	197	21	AAV08179	Arabidopsis thalia
144	34	81.0	197	21	AAV52233	Arabidopsis thalia
145	34	81.0	204	21	AAV58833	Breast and ovarian
146	34	81.0	224	22	AAV16512	Human novel secret
147	34	81.0	227	11	AAV05544	PesOD (superoxide
148	34	81.0	227	11	AAV05827	Iron superoxide di
149	34	81.0	236	21	AAV81244	Human Cdp (Cdc42hs
150	34	81.0	236	21	AAV81250	Human Cdp (Cdc42hs

ALIGNMENTS

ID	AA672246 standard; peptide: 7 AA.	XX
XX		XX
AC	AA672246;	AC
XX		XX
DT	14-MAY-2001 (first entry)	DT
XX		XX
DE	Colostrinin derived cytokine inducing peptide SEQ ID 1.	DE
XX		XX
KW	Colostrinin; immune response; cytokine; blood cell proliferation;	KW
KW	central nervous system disorder; neurological disorder; mental disorder;	KW
KW	dementia; neurodegenerative disease; Alzheimer's disease; psychosis;	KW
KW	neurosis; infection.	KW
OS	Synthetic.	OS
XX		XX
PM	WO200111937-A2.	PM
XX		XX
PD	22-FEB-2001.	PD
XX		XX
PF	17-AUG-2000; 2000WO-US22818.	PF
XX		XX
PR	17-AUG-1999; 99US-0149311.	PR
XX		XX
PA	(TEXA) UNIV TEXAS SYSTEM.	PA
PA	(REGE-) REGEN THERAPEUTICS PLC.	PA
XX		XX
PI	Stanton GJ, Hughes TK, Boldogh I, Georgiades J;	PI
XX		XX
DR	WPI: 2001-202804/20.	DR
XX		XX
PT	Inducing a cytokine and modulating an immune response; useful for	PT
PT	treating central nervous system diseases and bacterial and viral	PT
PT	infections; comprises administering colostrinin as an immunological	PT
PT	regulator -	PT
XX		XX
PS	Claim 1; Page 34; 50pp; English.	PS
XX		XX
CC	Sequences AAB72246 - AAB72275 represent peptides derived from colostrinin,	CC
CC	a proline rich polypeptide aggregate contained in colostrinum. The	CC
CC	peptides have immune response modulatory activity, and are capable of	CC
CC	inducing cytokines. Colostrinin and its derived peptides are useful for	CC
CC	inducing cytokine production, for modulating an immunological response	CC
CC	and for inducing blood cell proliferation. The peptides are useful in the	CC
CC	treatment of disorders of the central nervous system, neurological	CC
CC	disorders, mental disorders, dementia, neurodegenerative diseases,	CC
CC	Alzheimer's disease, motor neurone disease, psychosis, neurosis, chronic	CC
CC	disorders of the immune system, bacterial and viral infections and	CC
CC	acquired immunological deficiencies.	CC
XX		XX
SO	Sequence 7 AA;	SO
XX		XX
Query Match	100.0%; Score 42; DB 22; Length 7;	
Best local Similarity	100.0%; Pred. No. 6.4e+05;	
Matches 7; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 MOPPLP 7	
DB	1 mppplp 7	
RESULT 2		
AA672500	standard; peptide: 7 AA.	
AC	AA672500;	
XX		
DT	09-MAY-2001 (first entry)	
XX		
DE	Colostrinin peptide #1.	
KW	Dermatological; oxidative stress regulator; colostrinin.	

OY 1 MOPPLP 7
 |||||
 DB 1 mpppplp 7

RESULT 5

ABBS9354
 ID AAB59354 standard; Peptide: 8 AA.

AC AAB59354;

DT 21-MAR-2001 (first entry)

DE Ewe colostrinin peptide fragment derived sequence #14.

XX Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;
 KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

OS Ovis sp.

XX WO200075173-A2.

PN 14-DEC-2000.

PE 02-JUN-2000; 2000WO-GB02128.

FR 02-JUN-1999; 99GB-0012852.

PA (REGG-) REGEN THERAPEUTICS PLC.

XX Georgiades JA;

PI WPI: 2001-071058/08.

XX Peptides having an N-terminal amino acid sequence isolated from
 PT colostrinin for treating e.g. disorders of the central nervous system
 PT and immune system, viral and bacterial infections, and diseases
 PT characterized by amyloid plaques -

XX Claim 8; Page 27; 63pp; English.

XX The present invention provides the sequences of a number of peptides
 CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide
 CC fragment of colostrum. These peptides can be used in the treatment of
 CC central nervous system disorders such as senile dementia, Parkinson's
 CC disease, Alzheimer's disease, psychosis and neurosis, immune system
 CC disorders such as bacterial and viral infections, to improve the
 CC development of a child's immune system, as a dietary supplement, and to
 CC promote the dissolution of beta-amyloid plaques.

SQ Sequence 8 AA;

Query Match 100.0%; Score 42; DB 22; Length 8;

Best Local Similarity 100.0%; Pred. No. 6.4e+05; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 |||||
 DB 2 mpppplp 8

RESULT 6

ABBS7510
 ID ABB67510 standard; Protein: 316 AA.

AC ABB67510;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 29322.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.
 XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EM;

XX WPI: 2001-656860/75.

XX N-PSDB; ABL11613.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Disclosure; SEQ ID NO 29322; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).

XX The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 316 AA;

Query Match 100.0%; Score 42; DB 22; Length 316;

Best Local Similarity 100.0%; Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 |||||
 DB 249 mpppplp 255

RESULT 7

AAU45697
 ID AAU45697 standard; Protein: 95 AA.

AC AAU45697;

DT 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #593.

XX SAPHO syndrome; synovitis; acne; pustulosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

PR	21-APR-2000; 2000US-199047P.	
PR	02-JUN-2000; 2000US-208841P.	
PR	07-JUL-2000; 2000US-216747P.	
XX		
PA	(CORI-) CORIXA CORP.	
XX		
XX	Skelly YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;	
PI	L'valsoneuve J, Zhang Y, Jen S, Carter D;	
DR	WPI; 2001-616774/71.	
DR	N-PSDB; AAS59526.	
XX		
PT	Propionibacterium acnes polypeptides and nucleic acids useful for	
PT	vaccinating against and diagnosing infections, especially useful for	
PT	treating acne vulgaris -	
XX		
PS	Claim 6; SEQ ID NO 6892; 1069pp; English.	
XX		
CC	Sequences AA039105-AA068017 represent Propionibacterium acnes immunogenic	
CC	polypeptides. The proteins and their associated DNA sequences are used in	
CC	the treatment, prevention and diagnosis of medical conditions caused by	
CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,	
CC	psoriasis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.	
CC	P. acnes is also involved in infections of bone, joints and the central	
CC	nervous system, however it is particularly involved in the inflammatory	
CC	lesions associated with acne vulgaris. A method for detecting the	
CC	presence or absence of P. acnes in a patient comprises contacting a	
CC	sample with a binding agent that binds to the proteins of the invention	
CC	and determining the amount of bound protein in the sample. The	
CC	polypeptides may be used as antigens in the production of antibodies	
CC	specific for P. acnes proteins. These antibodies can be used to	
CC	deregulate expression and activity of P. acnes polypeptides and	
CC	therefore treat P. acnes infections. The antibodies may also be used as	
CC	diagnostic agents for determining P. acnes presence, for example, by	
CC	enzyme linked immunosorbent assay (ELISA).	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX		
XX	Sequence 95 AA;	
XX		
SO		
Query Match	92.98;	Score 39; DB 22; Length 95;
Best Local Similarity	85.7%;	Pred. NO. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;		
OY	1 MOPPLP 7	
	1:	
Db	88 mepplp 94	
RESULT 8		
ABB31347		
ID	ABB31347 standard; Peptide; 132 AA.	
XX		
AC	ABB31347;	
XX		
DT	01-FEB-2002 (first entry)	
XX		
DE	Peptide #3998 encoded by breast cell single exon nucleic acid probe.	
XX		
KW	Human; microarray; single exon probe; gene expression; breast;	
XX	disease; cancer.	
OS	Homo sapiens.	
XX		
PN	WO200157271-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	30-JAN-2001; 2001WO-US00662.	
XX		
PR	04-FEB-2000; 2000US-0180312.	

PR	26-MAY-2000;	2000US-0207456.
PR	30-JUN-2000;	2000US-0608408.
PR	03-AUG-2000;	2000US-0632366.
PR	21-SEP-2000;	2000US-0234687.
PR	27-SEP-2000;	2000US-0236359.
PR	04-OCT-2000;	2000GB-0024263.
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX		
DR	WPI; 2001-496933/54.	
XX		
PT	New spatially-addressable set of single exon nucleic acid probes,	
PT	useful for measuring gene expression in sample derived from human	
PT	breast, comprises number of single exon nucleic acid probes	-
XX		
PS	Claim 27; SEQ ID NO 14315; 327bp + sequence listing; English.	
CC	The invention relates to a spatially-addressable set of single exon	
CC	nucleic acid probes for measuring gene expression in a sample derived	
CC	from human breast and BT 474 cells. The method involves contacting	
CC	the probes with a collection of detectably labelled nucleic acids	
CC	derived from mRNA of human breast, and then measuring the label	
CC	bound to each probe of the microarray. The probes are useful for	
CC	verifying the expression of regions of genomic DNA predicted to	
CC	encode proteins. They are useful for gene discovery, and for	
CC	determining predisposition and/or prognosing breast disease. Gene	
CC	expression analysis is useful for assessing the toxicity of chemical	
CC	agents on cells. The microarray of this invention presents a far greater	
CC	diversity of probes for measuring gene expression, with far less bias	
CC	than expressed sequence tag microarrays. The method is suitable for	
CC	rapid production of functional information from genomic sequence. The	
CC	present sequence is a peptide encoded by a single exon nucleic acid	
CC	probe of the invention.	
CC	Note: The sequence data for this patent did not form part of the	
CC	printed specification, but was obtained in electronic format directly	
CC	from WIPO at ftp.wipo.int/pub/published_pat_sequences.	
XX		
SQ	Sequence 132 AA;	
DB	Query Match	92.9%; Score 39; DB 22; Length 132;
OY	Best Local Similarity	85.7%; Pred. No. 1.4e+02;
Matches	6; Conservative 1; Mismatches 0; Indels 0; Gaps	0;
DB	1 MQPPPLP 7 : 116 lqppplp 122	
RESULT	9	
ID	ABB36552	ABB36552 standard; Peptide; 132 AA.
AC	ABB36552;	
DT	04-FEB-2002 (first entry)	
DE	Peptide #4058 encoded by human foetal liver single exon probe.	
KW	Human; foetal liver; gene expression; single exon nucleic acid probe.	
OS	Homo sapiens.	
PN	WO200157277-A2.	
PD	09-AUG-2001.	
PF	30-JAN-2001; 2001WO-US00669.	
XX		
PR	04-FEB-2000; 2000US-0180312.	
PR	26-MAY-2000; 2000US-0207456.	

PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX
XX Claim 27; SEQ ID NO 29187; 639pp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a peptide encoded by a single exon
XX nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 132 AA:
SQ
OY 1 MOPPELP 7
DB 116 lqpplp 122
Query Match 92.9%; Score 39; DB 22; Length 132;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
RESULT 10
ABB21889
ID ABB21889 standard; Protein; 132 AA.
AC ABB21889;
XX
XX 23-JAN-2002 (first entry)
XX
XX Protein #3888 encoded by probe for measuring heart cell gene expression.
XX
XX Human; gene expression; heart; microarray; vascular system;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease.
XX
XX Homo sapiens.
XX
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001MO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX

DR WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
XX Claim 15; SEQ ID NO 23659; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX ABA21535-ABA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarray.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 132 AA:
SQ
OY 1 MOPPELP 7
DB 116 lqpplp 122
Query Match 92.9%; Score 39; DB 22; Length 132;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
RESULT 11
AAM57317
ID AAM57317 standard; Protein; 132 AA.
AC AAM57317;
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 29422.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001MO-US00667.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 29422; 650pp + Sequence Listing; English.
XX
XX

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.

XX
 SQ Sequence 132 AA;

Query Match 92.9%; Score 39; DB 22; Length 132;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 :|||||
 Db 116 lqpplp 122

RESULT 12

AAM69716
 ID AAM69716 standard; Protein; 132 AA.

XX
 AC AAM69716;

XX
 DT 06-NOV-2001 (first entry)

XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 30022.

XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 microarray; cancer; leukemia; lymphoma; myeloma.

XX
 OS Homo sapiens.

XX
 PN WO200157276-A2.

XX
 PD 09-AUG-2001.

XX
 PF 30-JAN-2001; 2001WO-US00668.

XX
 PR 04-FEB-2000; 2000US-0180312.

XX
 PR 26-MAY-2000; 2000US-0207456.

XX
 PR 30-JUN-2000; 2000US-0608408.

XX
 PR 03-AUG-2000; 2000US-0632366.

XX
 PR 21-SEP-2000; 2000US-0234687.

XX
 PR 27-SEP-2000; 2000US-0236359.

XX
 PR 04-OCT-2000; 2000GB-0024263.

XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
 DR WPI; 2001-488900/53.

XX
 PT Human genome-derived single exon nucleic acid probes useful for

XX
 PT analyzing gene expression in human bone marrow -

XX
 PS Example 4; SEQ ID NO: 30022; 658bp + Sequence Listing; English.

XX
 SQ Sequence 132 AA;

Query Match 92.9%; Score 39; DB 22; Length 132;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 :|||||
 Db 116 lqpplp 122

RESULT 13

AAM17529
 ID AAM17529 standard; Protein; 132 AA.

XX
 AC AAM17529;

XX
 DT 12-OCT-2001 (first entry)

XX
 DE Peptide #3963 encoded by probe for measuring cervical gene expression.

XX
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 cervical cancer.

XX
 OS Homo sapiens.

XX
 PN WO200157278-A2.

XX
 PD 09-AUG-2001.

XX
 PF 30-JAN-2001; 2001WO-US00670.

XX
 PR 04-FEB-2000; 2000US-0180312.

XX
 PR 26-MAY-2000; 2000US-0207456.

XX
 PR 30-JUN-2000; 2000US-0608408.

XX
 PR 03-AUG-2000; 2000US-0632366.

XX
 PR 21-SEP-2000; 2000US-0234687.

XX
 PR 27-SEP-2000; 2000US-0236359.

XX
 PR 04-OCT-2000; 2000GB-0024263.

XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
 DR WPI; 2001-488901/53.

XX
 PT Human genome-derived single exon nucleic acid probes useful for

XX
 PT analyzing gene expression in human cervical epithelial cells -

XX
 PS Claim 27; SEQ ID NO 22355; 487bp; English.

XX
 CC The present invention relates to human single exon nucleic acid probes

XX
 CC (SEMP: see A110068-A138459). The present sequence is a peptide encoded

XX
 CC by one such probe. The SENPs are derived from human Hela cells. The SENPs

XX
 CC can be used to produce a single exon microarray, which can be used for

XX
 CC measuring human gene expression in a sample derived from human cervical

XX
 CC epithelial cells. By measuring gene expression, the probes are therefore

XX
 CC useful in grading and/or staging of diseases of the cervix, notably

XX
 CC cervical cancer.

XX
 CC Note: The sequence data for this patent did not form part of the printed

XX
 CC specification, but was obtained in electronic format directly from WPIO

XX
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 132 AA;

Query Match 92.9%; Score 39; DB 22; Length 132;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 :|||||
 Db 116 lqpplp 122

RESULT 14

AAM30051
 ID AAM30051 standard; Protein; 132 AA.

```

XX AA030051;
AC 17-OCT-2001 (first entry)
XX
DT Peptide #4088 encoded by probe for measuring placental gene expression.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
KM genetic disorder.
XX
OS Homo sapiens.
XX
PN M0200157272-A2.
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001MO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOL-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 27; SEQ ID NO 30320; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI3115-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 132 AA:

Query Match          92.9%; Score 39; DB 22; Length 132;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0.

OY      1 MOPPLP 7
       :|11111
Db      116 lpppip 122

RESULT 15
AA05201
ID AA05201 standard; Protein; 132 AA.
XX
AC AA05201;
XX
DT 09-OCT-2001 (first entry)
DE Peptide #3883 encoded by probe for measuring breast gene expression.
XX
XX Probe; human; breast disease; breast cancer; development disorder;
KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN M0200157270-A2.
XX
```

```

XX 09-AUG-2001.
PD
PE 29-JAN-2001; 2001WO-US00661.
PR
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
PA
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
DR WPI; 2001-476286/51.
PT Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast -
PS Claim 27; SEQ ID No 13941; 322pp; English.
XX
XX The present invention relates to novel single exon nucleic acid probes
CC (see AI00010-AI110067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 132 AA:
SQ
Query Match 92.9%; Score 39; DB 22; Length 132;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0.
QY 1 MQPPPLP 7
DB 116 lqppplp 122
:::|||||
RESULT 16
AAE10123
ID AAE10123 standard; Protein; 38 AA.
XX
XX AAE10123;
AC
XX
DT 29-NOV-2001 (first entry)
XX
XX Human Ion channel-52 (ion52) protein.
DE
XX
XX Human: ion channel-52; ion52; antiinflammatory; immunosuppressive;
KW analgesic; nootropic; neuroprotective; antidepressant; cardiac;
KW cytosolic; antiviral; human immunodeficiency virus; HIV; anorectic;
KW antiviral; thyroid disorder; thyrotoxicosis; myxoedema; renal failure;
KW Crohn's disease; Rheumatoid arthritis; autoimmune disorder; pain;
KW stroke; psychotic disorder; neurological disorder; anxiety; dyskinesia;
KW Huntington's disease; degenerative disorder; Parkinson's disease;
KW schizophrenia; Alzheimer's disease; cardiovascular disease; cancer;
KW metabolic disorder; anorexia; obesity; mental disorder.
OS Homo sapiens.
XX
XX
PN WO200168849-A2.
XX
```

PD 20-SEP-2001.
 XX
 PF 09-MAR-2001; 2001WO-US07503.
 XX
 PR 10-MAR-2000; 2000US-0188400.
 PR 10-MAR-2000; 2000US-0188517.
 PR 10-MAR-2000; 2000US-0188518.
 PR 10-MAR-2000; 2000US-0188519.
 PR 05-JUL-2000; 2000US-0216815.
 PR 06-JUL-2000; 2000US-0216481.
 XX
 PA (PNA) PHARMACIA & UPJOHN CO.
 PI Wood LS, Vogel G, Karnovsky AM, Ruble CL, Linske-O'Connell LI;
 PI Wang J, Liu D;
 DR WPI; 2001-565795/63.
 DR N-PSDB; AAD17175.
 XX
 PT New ion channel polynucleotides and polypeptides, useful for
 PT identification of ion channel modulators and treatment of mental
 PT disorders, infections, cancer and autoimmune diseases.
 XX
 PS Claim 31; Page 102; 188pp; English.
 XX
 CC The patent discloses novel human ion channel polypeptides and their
 CC corresponding polynucleotides. The ion channel sequences and their
 CC modulators are used for the treatment of viral infections (e.g. human
 CC immunodeficiency virus (HIV)), thyroid disorders (e.g., thyrotoxicosis,
 CC myxoedema), renal failure, inflammatory conditions (e.g., Crohn's
 CC disease), rheumatoid arthritis, autoimmune disorders, pain, stroke,
 CC psychotic and neurological disorders (e.g. anxiety, depression and
 CC schizophrenia), dyskinesias (e.g. Huntington's disease), degenerative
 CC disorders (e.g., Parkinson's disease, Alzheimer's disease), cardio-
 CC vascular diseases, cancer, metabolic disorders (e.g. anorexia, obesity)
 CC and mental disorders. The present sequence is ion channel-52 (ion52)
 CC protein from human.
 XX
 SQ Sequence 38 AA;

Query Match 88.1%; Score 37; DB 22; Length 38;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 qpplp 7
 |||||
 DB 25 qpplp 30

RESULT 17
 AAB41170
 ID AAB41170 standard; Protein; 73 AA.
 XX
 AC AAB41170;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF934 polypeptide sequence SEQ ID NO:1868.
 XX
 KW Human: open reading frame; ORFX; detection; cytosolic; hepatotropic;
 KW vulnery; antipapillary; antipapillary; antipapillary; antipapillary;
 KW anticonvulsant; osteopathic; antipapillary; antipapillary; antipapillary;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antineoplastic;
 KW antiviral; antibacterial; antifungal; antineoplastic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antineoplastic disease; coagulation;

KM thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA) CURAGEN CORP.
 PI Shinkets RA, Leach M;
 PI WPI; 2000-602362/57.
 DR N-PSDB; AAC75379.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 1439; 5507pp; English.
 XX
 CC AAC74446 to AAC7606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytosolic; hepatotropic; vulnery;
 CC antipapillary; antipapillary; antipapillary; antipapillary;
 CC osteopathic; anticonvulsant; antipapillary; immunosuppressive;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antineoplastic; antibacterial; antiviral; antifungal; antineoplastic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antineoplastic disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 73 AA;

Query Match 88.1%; Score 37; DB 21; Length 73;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 qpplp 7
 |||||
 DB 55 qpplp 60

RESULT 18
 ABG21168
 ID ABG21168 standard; Protein; 95 AA.
 XX
 AC ABG21168;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #21159.
 XX
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;

KM food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX 31-MAR-2000; 2000US-0540217.
PR
XX 23-AUG-2000; 2000US-0649167.
PA
XX (HYSE-) HYSEQ INC.
PI
XX Drmanac RT, Liu C, Tang YT;
DR
XX WPI; 2001-639362/73.
DR
XX N-PSDB; AAS65355.
PT
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
PS
XX Claim 20; SEQ ID No 51527; 103pp; English.
PS
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AB00010-AB0377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 95 AA;
SO

Query Match 88.1%; Score 37; DB 22; Length 95;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 QPPPLP 7
Db 85 qppplp 90
IIIIII
RESULT 19
AAG3978
ID AAG3978 standard; Protein; 99 AA.
XX
XX AAG3978;
AC
XX
XX 17-OCT-2000 (first entry)
DE
XX Arabidopsis thaliana protein fragment SEQ ID NO: 27481.
KM
XX Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter.

KM termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
PN
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR
XX 05-MAR-1999; 99US-0123180.
PR
XX 09-MAR-1999; 99US-0123548.
PR
XX 23-MAR-1999; 99US-0125788.
PR
XX 25-MAR-1999; 99US-0126264.
PR
XX 29-MAR-1999; 99US-0126785.
PR
XX 01-APR-1999; 99US-0127462.
PR
XX 06-APR-1999; 99US-0128234.
PR
XX 08-APR-1999; 99US-0128714.
PR
XX 16-APR-1999; 99US-0129845.
PR
XX 19-APR-1999; 99US-0130077.
PR
XX 21-APR-1999; 99US-0130449.
PR
XX 23-APR-1999; 99US-0130510.
PR
XX 23-APR-1999; 99US-0130891.
PR
XX 28-APR-1999; 99US-0131449.
PR
XX 30-APR-1999; 99US-0132048.
PR
XX 30-APR-1999; 99US-0132407.
PR
XX 04-MAY-1999; 99US-0132484.
PR
XX 05-MAY-1999; 99US-0132485.
PR
XX 06-MAY-1999; 99US-0132486.
PR
XX 06-MAY-1999; 99US-0132487.
PR
XX 07-MAY-1999; 99US-0132863.
PR
XX 11-MAY-1999; 99US-0134256.
PR
XX 14-MAY-1999; 99US-0134218.
PR
XX 14-MAY-1999; 99US-0134221.
PR
XX 14-MAY-1999; 99US-0134370.
PR
XX 18-MAY-1999; 99US-0134768.
PR
XX 19-MAY-1999; 99US-0134941.
PR
XX 20-MAY-1999; 99US-0135124.
PR
XX 21-MAY-1999; 99US-0135353.
PR
XX 24-MAY-1999; 99US-0135629.
PR
XX 25-MAY-1999; 99US-0136021.
PR
XX 27-MAY-1999; 99US-0136392.
PR
XX 28-MAY-1999; 99US-0136782.
PR
XX 01-JUN-1999; 99US-0137222.
PR
XX 03-JUN-1999; 99US-0137528.
PR
XX 04-JUN-1999; 99US-0137502.
PR
XX 07-JUN-1999; 99US-0137724.
PR
XX 08-JUN-1999; 99US-0138094.
PR
XX 10-JUN-1999; 99US-0138540.
PR
XX 10-JUN-1999; 99US-0138847.
PR
XX 14-JUN-1999; 99US-0139119.
PR
XX 16-JUN-1999; 99US-0139452.
PR
XX 16-JUN-1999; 99US-0139453.
PR
XX 17-JUN-1999; 99US-0139452.
PR
XX 18-JUN-1999; 99US-0139454.
PR
XX 18-JUN-1999; 99US-0139455.
PR
XX 18-JUN-1999; 99US-0139456.
PR
XX 18-JUN-1999; 99US-0139457.
PR
XX 18-JUN-1999; 99US-0139458.
PR
XX 18-JUN-1999; 99US-0139459.
PR
XX 18-JUN-1999; 99US-0139460.
PR
XX 18-JUN-1999; 99US-0139461.
PR
XX 18-JUN-1999; 99US-0139462.
PR
XX 18-JUN-1999; 99US-0139463.
PR
XX 18-JUN-1999; 99US-0139750.
PR
XX 18-JUN-1999; 99US-0139763.
PR
XX 21-JUN-1999; 99US-0139817.
PR
XX 22-JUN-1999; 99US-0139899.
PR
XX 23-JUN-1999; 99US-0140353.
PR
XX 23-JUN-1999; 99US-0140354.
PR
XX 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0150885.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155148.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161459.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.1%; Score 37; DB 21; Length 99;
Best Local Similarity 100.0%; Pred.No.2,1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPLP 7
| | | | |
Db 81 qpplp 86

RESULT 20
AAR79095
ID AAR79095 standard; Protein: 278 AA.
XX AAR79095;
AC
XX
XX 20-FEB-1996 (first entry)
DT
XX
XX
DE Rat Fas ligand encoded by cDNA clone pTN24-15.
XX
XX
KW Fas ligand; Tumour Necrosis factor family; apoptosis; cell death;
KW Fas cell surface antigen; rat; Fas-L.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
FT Region 77..99
FT /label= transmembrane_anchor
FT Modified-site 116

```

FT Modified-site /label- N-glycosylation_site
FT 130
FT Modified-site /label- N-glycosylation_site
FT 247
FT Modified-site /label- N-glycosylation_site
FT 257
FT Modified-site /label- N-glycosylation_site
FT 257
XX MO9513293-A1.
XX
XX 18-MAY-1995.
XX
XX 10-NOV-1994; 94WO-JP01899.
XX
XX 18-OCT-1994; 94JP-0278378.
XX 10-NOV-1993; 93JP-0305975.
XX 13-DEC-1993; 93JP-0342526.
XX 18-MAR-1994; 94JP-0074344.
XX 08-JUL-1994; 94JP-0180955.
XX 07-SEP-1994; 94JP-0239363.
XX
XX (MOCH ) MOCHIDA PHARM CO LTD.
XX (OSAB-) OSAKA BIOSCIENCE INST.
XX
XX Nagata S, Nakamura N, Suda T, Takahashi T;
XX
XX WPI: 1995-194031/25.
XX N-PSDB: AA094153, AA03656.
XX
XX Peptide which binds to Fas antigen, and antibody reactive with it
XX for treatment and diagnosis of viral or auto-immune diseases
XX
XX Claim 8; Fig 2 and 3; 300pp: Japanese.
XX
XX Clones contg. rat cDNA fragments that encoded Fas ligand-like
XX proteins were identified by antibody panning. Clone pTN24-15 was
XX isolated and sequenced (AA094153). The sequence includes an open
XX reading frame coding for a predicted amino acid sequence of
XX 278 residues. The deduced protein sequence (AA094153) contains a
XX transmembrane anchor region and 4 consensus N-glycosylation sites;
XX the protein is a claimed Fas ligand able to induce apoptosis in
XX cells which express the Fas cell surface antigen.
XX
XX Sequence 278 AA:
SQ
Query Match 88.1%; Score 37; DB 16; Length 278;
Best Local Similarity 100.0%; Pred. No. Se+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 OPPPLP 7
DB 60 qppp 65
RESULT 21
AAW98069
ID AAW98069 standard; Protein; 278 AA.
XX
XX AAW98069;
XX
XX 21-JUN-1999 (first entry)
XX
XX Rat Fas ligand (FasL).
XX
XX FasL; Fas ligand; proinflammatory; immunosuppressive; rat;
XX graft versus host disease; autoimmune disease; psoriasis;
XX rheumatoid arthritis; systemic lupus erythematosus; gene therapy.
XX
XX Rattus sp.
XX
XX Key Location/Qualifiers
XX PH Domain 78..99

```

```

FT Modified-site /note- "transmembrane domain"
FT 116
FT Modified-site /note- "N-glycosylated"
FT 130
FT Modified-site /note- "N-glycosylated"
FT 247
FT Modified-site /note- "N-glycosylated"
FT 257
FT Modified-site /note- "N-glycosylated"
FT 257
XX MO9903999-A1.
XX
XX 28-JAN-1999.
XX
XX 16-JUL-1998; 98WO-US14771.
XX 17-JUL-1997; 97US-0052829.
XX
XX (UNMI ) UNIV MICHIGAN.
XX
XX Chen J, Nabel GJ;
XX
XX WPI: 1999-132243/11.
XX N-PSDB: AAX24877.
XX
XX Inhibition of proinflammatory responses - using an agent which
XX modulates FasL stimulation, used for treating graft versus host
XX disease or autoimmune disease
XX
XX Disclosure; Fig 4A: 71pp: English.
XX
XX This present sequence is rat Fas ligand (FasL). The invention
XX provides a method for inhibiting a proinflammatory response in a
XX cell mixture by administering an immunosuppressive agent which
XX inhibits the proinflammatory activity of FasL. In some embodiments,
XX FasL is coadministered with the immunosuppressive agent, and the
XX cell mixture comprises neutrophil cells. The method can be
XX practised in vitro, ex vivo or in vivo. Suitable immunosuppressive
XX agents include antisense molecules that inhibit endogenous FasL
XX expression, soluble Fas receptors or variants, ribozymes that
XX inhibit the endogenous expression of FasL, drugs that inhibit FasL
XX signalling, agents that induce the endogenous expression of FasL
XX transforming growth factor (TGF)-beta, and polynucleotides coding
XX for an immunosuppressive agent such as TGF-beta. The method can be
XX used for treating diseases associated with an undesired FasL-mediated
XX proinflammatory response, e.g. graft versus host disease, or an
XX autoimmune disease such as systemic lupus erythematosus, rheumatoid
XX arthritis and psoriasis. The invention also provides a method for
XX identifying agents which modulate FasL stimulation of a
XX proinflammatory response.
XX
XX Sequence 278 AA:
SQ
Query Match 88.1%; Score 37; DB 20; Length 278;
Best Local Similarity 100.0%; Pred. No. Se+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 OPPPLP 7
DB 60 qppp 65
RESULT 22
AAW95040
ID AAW95040 standard; Protein; 278 AA.
XX
XX AAW95040;
XX
XX 14-MAY-1999 (first entry)
XX
XX Rat FasL protein.
XX

```


KM Cell proliferation; Fasl protein; pathogen; modulation; cell locus;
KM Proinflammatory response; inhibition; rat.
XX Rattus sp.
XX WO9903998-A1.
XX
XX 28-JAN-1999.
XX
XX 16-JUL-1998; 98WO-0514770.
XX
XX 17-JUL-1997; 97US-0052829.
XX
XX (UNMI) UNIV MICHIGAN.
XX
XX Nabel GJ;
XX
XX WPI: 1999-132242/11.
XX N-PSDB: AAX21573.
XX
XX Method for inhibiting proliferation of a cell - comprises
XX administration of Fasl protein to the cell locus
XX
XX Disclosure: Fig 6A-C; 70pp; English.
XX
XX The invention relates to a method for inhibiting the proliferation of a
XX suitable cell or pathogen which comprises administering to the cells
XX locus an effective amount of Fasl. A method for identifying agents which
XX modulate Fasl stimulation of a localised proinflammatory response is also
XX provided. The methods can be used for provoking a proinflammatory
XX response in the locus of a cell. It can also be used to inhibit the
XX proliferation of a cell or a pathogen in a subject. The present sequence
XX represents a rat Fasl protein.
XX
SQ Sequence 278 AA:

Query Match 88.1%; Score 37; DB 20; Length 278;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
DB 60 qppplp 65

RESULT 23

AAI50933
ID AAI50933 standard; Protein; 302 AA.

AC AAI50933;

DT 10-MAR-2000 (first entry)

DE Human fetal brain cDNA clone vc26_1 derived protein #2.

KM Human; secreted protein; treatment; nutritional activity; cytokine;
KM cell proliferation; cell differentiation; hematopoiesis regulation;
KM tissue growth; activin; inhibin; chemotactic; chemokine; hemostatic;
KM thrombolytic; anti-inflammatory; invasion suppressor; tumor inhibition;
KM gene therapy.
XX
XX Homo sapiens.
XX
XX WO9955721-A1.
XX
XX 04-NOV-1999.
XX
XX 23-APR-1999; 99WO-US08504.
XX
XX 24-APR-1998; 98US-0082804.
XX 11-JUN-1998; 98US-0088994.
XX 12-JUN-1998; 98US-0089278.

PR 02-JUL-1998; 98US-0091647.
PR 24-AUG-1998; 98US-0097639.
PR 22-APR-1999; 99US-0097639.
XX
XX (ALPH-) ALPHAGE INC.
XX
XX Valenzuela D, Yuan O, Hoffman H, Hall J, Ragle]ko P;
XX
XX WPI: 2000-052801/04.
XX N-PSDB: AA43798.
XX
XX New polynucleotides encoding secreted human proteins, derived from
XX human fetal brain, adult skin, adult brain, adult heart, adult thymus
XX and adult aorta cDNA libraries.
XX
XX Disclosure: Page 274-275; 282pp; English.
XX
XX This invention describes novel human secreted proteins which are encoded
XX by polynucleotides obtained from fetal brain, adult skin, adult brain,
XX adult heart, adult thymus and adult aorta cDNA libraries. The
XX polynucleotides and proteins are predicted to have biological activities
XX which would make them suitable for treating, preventing or ameliorating
XX medical conditions in humans and animals, although no supporting data
XX is given. Suggested activities include nutritional activity, cytokine
XX and cell proliferation/differentiation activity. Immune stimulating
XX (e.g. as vaccines) or suppressing activity, hematopoiesis regulating
XX activity, tissue growth activity, activin/inhibin activity,
XX chemotactic/chemokinetic activity, hemostatic and thrombolytic activity,
XX receptor/ligand activity, anti-inflammatory activity, cadherin/tumor
XX invasion suppressor activity, and tumor inhibition activity. The
XX polynucleotides are also stated to be useful for gene therapy.
XX AAI50905-Y50947 represent the secreted proteins described in the method
XX of the invention which are encoded by the polynucleotides represented in
XX AA43777-243808.
XX
SQ Sequence 302 AA:

Query Match 88.1%; Score 37; DB 21; Length 302;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
DB 249 qppplp 254

RESULT 24

AAE10122
ID AAE10122 standard; Protein; 332 AA.

AC AAE10122;

DT 29-NOV-2001 (first entry)

DE Human ion channel-31c4 (lon31c4) protein.

KM Human; ion channel-31c4; lon31c4; antiinflammatory; immunosuppressive;
KM analgesic; nootropic; neuroprotective; antidepressant; cardiant;
KM cytoskeletal; antiviral; human immunodeficiency virus; HIV; anorectic;
KM antiviral; thyroid disorder; thyrotoxicosis; myxoedema; renal failure;
KM Crohn's disease; rheumatoid arthritis; autoimmune disorder; pain;
KM stroke; psychotic disorder; neurological disorder; anxiety; dyskinesia;
KM Huntington's disease; degenerative disorder; Parkinson's disease;
KM schizophrenia; Alzheimer's disease; cardiovascular disease; cancer;
KM metabolic disorder; anorexia; obesity; mental disorder.
XX
XX Homo sapiens.
XX
XX WO200168849-A2.
XX
XX 20-SEP-2001.

PF 09-MAR-2001; 2001WO-US07503.
 XX
 PR 10-MAR-2000; 2000US-0188400.
 PR 10-MAR-2000; 2000US-0188517.
 PR 10-MAR-2000; 2000US-0188518.
 PR 10-MAR-2000; 2000US-0188519.
 PR 05-JUL-2000; 2000US-0216815.
 PR 06-JUL-2000; 2000US-0216841.
 XX
 PA (PHARMA) PHARMACIA & UPJOHN CO.
 XX
 PI Wood LS, Vogel G, Karnovsky AM, Ruble CL, Linske-O'Connell LI;
 PI Wang J, Liu D;
 XX
 DR WPI: 2001-565795/63.
 DR N-PSDB: AAD17174.
 XX
 PT New ion channel polynucleotides and polypeptides, useful for
 PT identification of ion channel modulators and treatment of mental
 PT disorders, infections, cancer and autoimmune diseases -
 XX
 PS Claim 90; Page 102; 186pp; English.
 XX
 CC The patent discloses novel human ion channel polypeptides and their
 CC corresponding polynucleotides. The ion channel sequences and their
 CC modulators are used for the treatment of viral infections (e.g. human
 CC immunodeficiency virus (HIV)), thyroid disorders (e.g., thyrotoxicosis,
 CC myxoedema), renal failure, inflammatory conditions (e.g., Crohn's
 CC disease), rheumatoid arthritis, autoimmune disorders, pain, stroke,
 CC psychotic and neurological disorders (e.g. anxiety, depression and
 CC schizophrenia), dyskinesias (e.g. Huntington's disease), degenerative
 CC disorders (e.g., Parkinson's disease, Alzheimer's disease), cardio-
 CC vascular diseases, cancer, metabolic disorders (e.g. anorexia, obesity)
 CC and mental disorders. The present sequence is ion channel-31c4 (ion31c4)
 CC protein from human.
 XX
 SQ Sequence 332 AA;
 XX
 Query Match 88.1%; Score 37; DB 22; Length 332;
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 OY 2 QPPLP 7
 DB 288 qpplp 293
 XX
 RESULT 25
 AAB83413
 ID AAB83413 standard; Protein; 393 AA.
 XX
 AC AAB83413;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human ion1 protein sequence #3.
 XX
 KW Human: ion2a; ion2b; ion3; ion4a; ion4b; ion5; ion6; ion7; Nootropic;
 KW cytostatic; immunosuppressive; neuroprotective; antiinflammatory;
 KW antithymatic; antiarthritic; antidiabetic; anorectic; virucide;
 KW anti-HIV; antiparkinsonian; antithyroid; hypotensive; hypertensive;
 KW anticonvulsant; tranquilizer; cerebroprotective; analgesic; anxiety;
 KW antipsoriatic; immune response; mental disorder; viral infection;
 KW thyroid disorder; renal failure; inflammatory condition; homeostasis;
 KW cell differentiation; rheumatoid arthritis; autoimmune disorder;
 KW movement disorder; CNS disorder; psychotic disorder; schizophrenia;
 KW neurological disorder; degenerative disorder; Parkinson's disease;
 KW Alzheimer's disease; metabolic disorder; cardiovascular disease;
 KW proliferative disease; cancer; hormonal disorder; sexual dysfunction;
 KW brain injury; therapy; ion1.
 XX
 OS Homo sapiens.

XX
 PN WO200144283-A2.
 XX
 PD 21-JUN-2001.
 XX
 PF 14-DEC-2000; 2000WO-US33829.
 XX
 PR 14-DEC-1999; 99US-0460602.
 XX
 PA (PHARMA) PHARMACIA & UPJOHN CO.
 XX
 PI Roberds SL, Karnovsky AM, Ruble CL, Benjamin CW;
 PI WPI: 2001-648142/74.
 XX
 DR Novel ion channel polynucleotides and polypeptides useful for
 DR identifying ion channel agonists/antagonists of therapeutic use and for
 DR diagnosing, treating mental disorders and metabolic diseases -
 XX
 PS Claim 35; Page 73; 157pp; English.
 XX
 CC This sequence is encoded by a human ion channel nucleic acid molecule of
 CC the invention. The invention relates to the human ion1, ion2a, ion2b,
 CC ion3, ion4a, ion4b, ion5, ion6, and ion7 proteins and their corresponding
 CC DNA sequences. The ion proteins of the invention have the following
 CC activities: Nootropic; cytostatic; immunosuppressive; neuroprotective;
 CC antiinflammatory; antithymatic; antiarthritic; antidiabetic; anorectic;
 CC virucide; anti-HIV; antiparkinsonian; antithyroid; hypotensive;
 CC hypertensive; anticonvulsant; tranquilizer; cerebroprotective; analgesic;
 CC antipsoriatic. The DNA sequences are useful for identifying a compound
 CC which binds a nucleic acid molecule encoding ion proteins using gel-shift
 CC assay. Ion proteins are useful for inducing an immune response and for
 CC identifying a compound which binds ion protein. The compounds identified
 CC as binding ion-1 or ion-3 are useful for treating mental disorders.
 CC Modulators of ion protein activity are useful for treating diseases and
 CC physiological conditions, such as viral infections caused by HIV-1,
 CC thyroid disorders, renal failure, inflammatory conditions, diseases
 CC related to cell differentiation and homeostasis, rheumatoid arthritis,
 CC autoimmune disorders, movement disorders, including ataxias, CNS
 CC disorders, psychotic and neurological disorders including anxiety,
 CC schizophrenia, degenerative disorders such as Parkinson's, Alzheimer's
 CC disease, metabolic and cardiovascular diseases and disorders,
 CC proliferative diseases and cancers, hormonal disorders and sexual
 CC dysfunction. Ion proteins are useful in treating acute and/or traumatic
 CC brain injury. Ion polynucleotides, polypeptides and modulators are also
 CC useful in diagnostic assays for such conditions or diseases. The proteins
 CC are useful as a diagnostic tool for disease or disorders and as research
 CC tools for identification, characterization and purification of
 CC interacting, regulatory proteins. Antibodies against the proteins are
 CC useful e.g. to treat neurological and psychiatric disorders.
 XX
 SQ Sequence 393 AA;
 XX
 Query Match 88.1%; Score 37; DB 22; Length 393;
 Best Local Similarity 100.0%; Pred. No. 6.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 OY 2 QPPLP 7
 DB 276 qpplp 281
 XX
 RESULT 26
 ID ABB60862 standard; Protein; 440 AA.
 XX
 AC ABB60862;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 9378.
 XX

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 OS Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-0509231.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PMD, Myers EW;
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL04955.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS Disclosure; SEQ ID NO 9378; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (AB057737-AB072072),
 CC (AB057737-AB072072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SO Sequence 440 AA;

Query Match 88.1%; Score 37; DB 22; Length 440;
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 OPPPLP 7
 Db 55 qppplp 60

RESULT 27
 AABG09852
 ID AABG09852 standard; Protein: 450 AA.
 XX
 AC AABG09852;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #9843.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS74039.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID NO 40211; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABB00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SO Sequence 450 AA;

Query Match 88.1%; Score 37; DB 22; Length 450;
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 OPPPLP 7
 Db 158 qppplp 163

RESULT 28
 AAB60493
 ID AAB60493 standard; Protein: 580 AA. -
 XX
 AC AAB60493;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Human cell cycle and proliferation protein CCYPR-41, SEQ ID NO:41.
 XX
 KW Cell cycle and proliferation protein; CCYPR; human; agonist;
 KW antagonist; gene therapy; detection; gene therapy;
 KW transgenic animal disease model; immune disorder;
 KW developmental disorder; cell signalling disorder;
 KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
 KW arteriosclerosis; asthma; allergy; diabetes mellitus;
 KW menstrual cycle disorder; bacterial infection.
 XX
 OS Homo sapiens.
 XX
 PN WO200107471-A2.
 XX
 PD 01-FEB-2001.

ID AAW48760 standard; Protein: 667 AA.
 XX
 AC AAW48760;
 XX
 DT 18-AUG-1998 (first entry)
 XX
 DE BOP1 protein.
 XX
 KW Corticotroph; pituitary; BOP1; tumour suppressor; constitutive;
 KW Inducible; Alzheimer's disease; nuclear transcription factor; apoptosis;
 KW cell cycle; neuronal disorder.
 XX
 OS Mus sp.
 XX
 PH Key
 FT Region
 FT /note= "contains 7 zinc finger motifs of the
 FT (Cys)2(His)2 type"
 FT Modified-site
 FT /note= 56.60
 FT "Cyclin dependant kinase consensus
 FT phosphorylation site"
 FT Modified-site
 FT /note= 666
 FT "Protein kinase A phosphorylation site"
 XX
 PN W09813489-A1.
 XX
 PD 02-APR-1998.
 XX
 PF 22-SEP-1997; 97WO-EP05198.
 XX
 PR 23-SEP-1996; 96US-0718661.
 XX
 PA (CNRS) CENT NAT RECH SCI.
 PA (PLAC) MAX PLANCK GES FORDERUNG WISSENSCHAFTEN.
 XX
 PI Journot L, Spengler D;
 XX
 DR WPI: 1998-230701/20.
 XX
 PT New isolated tumour suppressor gene - useful for developing products
 PT for use in diagnosis and treatment of tumour(s) or neuronal
 PT disorder(s)
 XX
 PS Claim 1; Fig 1; 118pp; English.
 XX
 CC The present sequence represents the BOP1 protein encoded by the
 CC BOP1 cDNA isolated from the mouse corticotroph pituitary tumour
 CC cell line Atr-20 cDNA library. The BOP1 protein displays a tumour
 CC suppressing activity when it was constitutively and inducibly
 CC expressed in tumour cells. The BOP1 cDNA and the protein it encodes
 CC are claimed to be useful in the preparation of therapeutic
 CC compositions, useful for treating, preventing or delaying the
 CC recurrence of a tumour or neuronal disorders, e.g. genetic diseases
 CC or acquired degenerative encephalopathies such as Alzheimer's disease.
 CC The BOP1 protein is also claimed to be able to induce apoptosis
 CC resulting in inhibition of tumour cell growth, to suppress tumour
 CC formation, to induce G1 arrest of the cell cycle and to act as nuclear
 CC transcription factor.
 XX
 SQ Sequence 667 AA;

Query Match 88.1%; Score 37; DB 19; Length 667;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 I I I I I I
 DB 527 qpplp 532

RESULT 31
 ABBG07912

ID ABBG07912 standard; Protein: 681 AA.
 XX
 AC ABBG07912;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #7903.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB; AAS72099.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 38271; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABBG00010-ABG30327 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 681 AA;

Query Match 88.1%; Score 37; DB 22; Length 681;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 I I I I I I
 DB 396 qpplp 401

RESULT 32
 ABB58741
 ID ABB58741 standard; Protein: 745 AA.

```

XX ABB58741;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster polypeptide SEQ ID NO 3015.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX
OS Drosophila melanogaster.
PN W0200171042-A2.
PD
XX 27-SEP-2001.
PP
XX 23-MAR-2001; 2001WO-US09231;
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
PA (PERE ) PE CORP NY.
PI Venter JC, Adams M, Li PWD, Myers EW;
DR WI: 2001-6556860/75.
DR N-PSDB; ABL02844.
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure: SEQ ID NO 3015; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-ABL30511), expressed DNA
CC sequences (AB101840-ABL16175) and the encoded proteins
CC (ABB57737-ABB12072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CX
XX
SQ Sequence 745 AA;

Query Match      88.1%; Score 37; DB 22; Length 745;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OPPPLP 7
   |||||
Db    655 qppplp 660

RESULT 33
ABB67658
ID ABB67658 standard; Protein; 819 AA.
AC ABB67658;
AD
DE 26-MAR-2002 (first entry)
DX
KW Drosophila melanogaster polypeptide SEQ ID NO 29766.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
OS Drosophila melanogaster.

```

PN WO200171042-A2.
PD
XX 27-SEP-2001.
PF 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, LI PWD, Myers EW;
DR N-PSDB; ABL11761.
XX WPI: 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PP genes from Drosophila and for elucidating cell signalling and cell-cell
PS interactions -
PS Disclosure: SEQ ID NO 29766; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-AB116175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 819 AA;

Query Match 88.1%; Score 37; DB 22; Length 819;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 2 OPPPLP 7
IIIIII
Db 125 gppplp 130

RESULT 34
AAW07600
ID AAW07600 standard; Protein; 874 AA.
XX
AC AAW07600;
XX
DT 26-FEB-1997 (first entry)
XX
DE Human G-protein receptor HCEGH45.
XX
KW G-protein receptor; HCEGH45; signal transduction;
KM pituitary adenylate cyclase activating polypeptide;
KW PACAP hypersecretion; Alzheimer's disease; Parkinson's disease;
KM amnesia; antagonist; agonist; diagnosis; therapy.
XX
OS Homo sapiens.
XX
PN MO9639439-A1.
XX
PD 12-DEC-1996.
XX
PF 06-JUN-1995; 95WO-US07188.
XX
PR 06-JUN-1995; 95WO-US07188.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX

PI L1 Y, Rosen CA, Ruben SM, Soppet DR;
 XX
 DR WPI: 1997-043074/04.
 XX
 PT Human G-protein pituitary adenylate cyclase activating polypeptide
 CC receptor, HCEGH45 - useful to identify (ant)agonists, for treatment
 CC of asthma, amnesia, Parkinson's disease, etc.
 PS
 CC Claim 1, Fig 1A-J, 65pp; English.
 CC
 CC Novel human mature G-protein receptor HCEGH45 (AA07600) is a
 CC G-protein pituitary adenylate cyclase activating polypeptide
 CC (PACAP)-like receptor of amnesia-like neuropeptides. It is a
 CC 7-transmembrane protein and participates in signal transduction
 CC pathways. Its amino acid sequence was deduced from a cDNA clone
 CC (AA0744039) derived from human cerebellum tissue. Recombinant HCEGH45
 CC can be produced in host (e.g. COS, Sf9) cells and used to identify
 CC (ant)agonists to the polypeptide. Antagonists can be used to treat
 CC PACAP hypersecretory conditions and to create amnesia models.
 CC Agonists are employed to treat amnesia and Alzheimer's disease.
 CC
 SO Sequence 874 AA:
 Query Match 88.1%; Score 37; DB 18; Length 874;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 OPPPLP 7
 DB 730 qppplp 735
 RESULT 35
 ID AAM94075 standard; Protein; 874 AA.
 XX
 AC AAM94075;
 XX
 DT 19-APR-1999 (first entry)
 XX
 DE Human G-protein coupled receptor polypeptide, HCEGH45.
 XX
 KW Human; G-protein coupled receptor; HCEGH45; PACAP; neuropeptide;
 KW pituitary adenylate cyclase activating polypeptide; hypersecretory;
 KW amnesia; Alzheimer's disease.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 708
 FT /label= unknown
 FT /note= "encoded by NNN"
 XX
 PN US5869632-A.
 XX
 PD 09-FEB-1999.
 XX
 PE 06-JUN-1995; 95US-0465976.
 XX
 PR 06-JUN-1995; 95OS-0465976.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI L1 Y, Rosen CA, Ruben SM, Soppet DR;
 XX
 DR WPI: 1999-152883/13.
 DR N-PSDB; AAX01094.
 XX
 PT DNA encoding human G-protein coupled receptor - useful for producing
 PT recombinant polypeptide
 XX
 PS Disclosure; Fig 1A-L; 37pp; English.

XX
 CC This represents a human G-protein coupled receptor polypeptide, HCEGH45.
 CC The encoding cDNA is deposited under the accession no: ATCC 97132. The
 CC polynucleotide is used to produce a recombinant HCEGH45 which has been
 CC putatively identified as a human G-protein pituitary adenylate cyclase
 CC activating polypeptide (PACAP)-like receptor for amnesia like
 CC neuropeptides. Antagonists of the polypeptide may be used therapeutically
 CC to treat PACAP hypersecretory conditions while agonists may be employed
 CC to treat amnesia and Alzheimer's disease.
 CC Note: this sequence (provided in the Figure 1A-L) differs from the
 CC HCEGH45 sequence given in the sequence listing after amino acid residue
 CC 687.
 CC
 SO Sequence 874 AA:
 Query Match 88.1%; Score 37; DB 20; Length 874;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 OPPPLP 7
 DB 730 qppplp 735
 RESULT 36
 ID AAB92916 standard; Protein; 906 AA.
 XX
 AC AAB92916;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:11554.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PE 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Salto K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI: 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 11554; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

CC Sequence 906 AA:

Query Match 88.1%; Score 37; DB 22; Length 906;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPLP 7
 |||||
 Db 81 qpplp 86

RESULT 37

AAH00834
 ID AAH00834 standard; Protein: 908 AA.

AC AAH00834;

DT 01-OCT-2001 (first entry)

DE Human bone marrow protein, SEQ ID NO: 197.

DE Human; bone marrow; antinflammatory; cytosolic; neuroprotective;
 KM antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
 KM immunosuppressive; gene therapy; cytokine cell proliferation;
 KM cell differentiation modulator; immune disorder; infection; cancer;
 KM human immunodeficiency virus; HIV; autoimmune disorder; hemophilia.

OS Homo sapiens.

PN WO200153453-A2.

PD 26-JUL-2001.

PF 23-DEC-2000; 2000MO-US34960.

XX 21-JAN-2000; 2000US-0468725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 30-NOV-2000; 2000US-0250583.

XX (HYSE-) HYSEQ INC.

PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
 PI Ren F, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Dimaacac RT;

XX WPI: 2001-488707/53.

DR N-PSDB; AAB89953.

PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
 PT for treating e.g. cancer and immune deficiency disorders -

PS Claim 10; Page 347-349; 648bp; English.

XX The present sequence is one of 251 novel human polypeptides encoded
 CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
 CC polypeptide encoded by it are useful in the treatment of various
 CC immune deficiencies and disorders. The deficiencies and disorders may
 CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
 CC infection, or may result from an autoimmune disorder, a coagulation
 CC disorder (e.g. hemophilia), inhibition of tumour cell proliferation,
 CC suppression of an inflammatory response or treatment of a nervous
 CC system disorder such as Alzheimer's disease. Detection of the presence
 CC or increased expression of the polynucleotide or the protein it
 CC encodes is useful for the diagnosis and/or prognosis of one
 CC or more types of cancer. The polynucleotide and polypeptide can be
 CC used as nutritional sources or supplements and in the screening of
 CC chemical compounds as potential drugs.

CC Sequence 908 AA:

Query Match 88.1%; Score 37; DB 22; Length 908;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPLP 7
 |||||
 Db 234 qpplp 239

RESULT 38

ABB59788
 ID ABB59788 standard; Protein: 961 AA.

AC ABB59788;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide seq ID NO 6156.

DE Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 KM Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001MO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EM;

XX WPI: 2001-656860/75.

DR N-PSDB; ABL03891.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 6156; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 961 AA:

Query Match 88.1%; Score 37; DB 22; Length 961;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
IIIIII
DB 725 QPPPLP 730

RESULT 39

AA079129
ID AA079129 standard; Protein; 1003 AA.

XX
AC AA079129;

XX
DT 06-NOV-2001 (first entry)

XX
DE Human protein SEQ ID NO 1791.

XX
DE Human: cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation.

XX
OS Homo sapiens.

XX
PN MO200157190-A2.

XX
PD 09-AUG-2001.

XX
PF 05-FEB-2001; 2001WO-US04098.

XX
PR 03-FEB-2000; 2000US-0496914.

XX
PR 27-APR-2000; 2000US-0560875.

XX
PR 20-JUN-2000; 2000US-0598075.

XX
PR 19-JUL-2000; 2000US-0620325.

XX
PR 01-SEP-2000; 2000US-0654936.

XX
PR 15-SEP-2000; 2000US-0663561.

XX
PR 30-OCT-2000; 2000US-0693325.

XX
PR 30-NOV-2000; 2000US-0728422.

XX
PA (HYSE-) HYSEQ INC.

XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;

XX
PI WPI; 2001-476283/51.

XX
DR N-PSDB; AAK52262.

XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,

XX
PT useful in diagnosis and gene therapy -

XX
PS Claim 20; Page 4150-4152; 6221pp; English.

XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

CC encoded polypeptides (AA0788323-AA080302) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AA080020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX
SQ Sequence 1003 AA:

Query Match 88.1%; Score 37; DB 22; Length 1003;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
IIIIII
DB 561 QPPPLP 566

RESULT 40

AA033177
ID AA033177 standard; Protein; 1217 AA.

XX
AC AA033177;

XX
DT 18-DEC-2001 (first entry)

XX
DE Novel human secreted protein #3668.

XX
DE Human: Vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX
OS Homo sapiens.

XX
PN MO200179449-A2.

XX
PD 25-OCT-2001.

XX
PF 16-APR-2001; 2001WO-US08656.

XX
PR 18-APR-2000; 2000US-0552929.

XX
PR 26-JAN-2001; 2001US-0770160.

XX
PA (HYSE-) HYSEQ INC.

XX
PI Tang YT, Liu C, Drmanac RT;

XX
PI WPI; 2001-611725/70.

XX
DR Nucleic acids encoding a range of human polypeptides, useful in genetic

XX
PT vaccination, testing and therapy -

XX
PS Claim 20; Page 719-720; 765pp; English.

XX
CC The invention relates to novel human-secreted polypeptides. The

XX
CC polypeptides and antibodies to the polypeptides are useful for

XX
CC determining the presence of or predisposition to a disease associated

XX
CC with altered levels of polypeptide. The polypeptides are also useful for

XX
CC identifying agents (agonists and antagonists) that bind to them. Cells

XX
CC expressing the proteins are useful for identifying a therapeutic agent

XX
CC for use in treatment of a pathology related to aberrant expression or

XX
CC physiological interactions of the polypeptide. Vectors comprising

XX
CC the nucleic acids encoding the polypeptides and cells genetically

XX
CC engineered to express them are also useful for producing the proteins.

XX
CC The proteins are useful in genetic vaccination, testing and

XX
CC therapy, and can be used as nutritional supplements. They may be used to

XX
CC increase stem cell proliferation; to regulate haematopoiesis; and in

XX
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;

XX
CC immune suppression and/or stimulation; as anti-inflammatory agents; and

XX
CC in treatment of leukaemias. AA029510-AA033304 represent the amino acid

XX
CC sequences of novel human secreted proteins of the invention.

XX
SQ Sequence 1217 AA:

Query Match 88.1%; Score 37; DB 22; Length 1217;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
111111
Db 234 qppplp 239

RESULT 41

AA067393
ID AA067393 standard; Protein; 1244 AA.

XX
AC AA067393;

DT 13-NOV-2001 (first entry)

DE Amino acid sequence of human protein kinase SGK223.

XX
KW Human; protein kinase; cancer; immune disease; cardiovascular disease;
KW brain disease; neuronal disease; Alzheimer's disease; chromosome 8;
KW Parkinson's disease; multiple sclerosis; metabolic disorder;
KW peripheral nervous system disease; amyotrophic lateral sclerosis;
KW infection; ocular disease; migraine; pain; sexual dysfunction;
KW mood disorder; attention disorder; cognition disorder; hypotension;
KW hypertension; psychotic disorder; dyskinesia; transplant rejection.

XX
OS Homo sapiens.

XX
PN WO200166594-A2.

XX
PD 13-SEP-2001.

XX
PP 02-MAR-2001; 2001WO-US06638.

XX
PR 06-MAR-2000; 2000US-0187150.

XX
PR 29-MAR-2000; 2000US-0193404.

XX
PR 13-NOV-2000; 2000US-0247013.

XX
PA (SUGEN) SUGEN INC.

XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;

XX
DR WPI; 2001-536777/59.

XX
DR N-PSDB; AA077992.

XX
PT Nucleic acids capable of encoding human polypeptides having a kinase or
PT kinase-like activity, useful for diagnosing a disease selected from
PT cancers, cardiovascular disease and neuronal-associated diseases (e.g.
PT Alzheimer's disease) -

XX
PS Claim 7; Fig 2A; 201pp; English.

XX
CC The present sequence represents a human protein kinase. The
CC gene is located at chromosomal position 8p22-p23. The kinase
CC polypeptides are useful for diagnosing a disease or disorder
CC selected from cancers (e.g. cancers of tissues and cancers of
CC hematopoietic origin), immune-related diseases and disorders,
CC cardiovascular disease, brain or neuronal-associated diseases (e.g.
CC Alzheimer's disease, Parkinson's disease, multiple sclerosis),
CC metabolic disorders, peripheral nervous system diseases, amyotrophic
CC lateral sclerosis, viral infections, infections caused by prions,
CC infections caused by bacteria, infections caused by fungi, ocular
CC diseases, migraines, pain, sexual dysfunction, mood disorders,
CC attention disorders, cognition disorders, hypotension, hypertension,
CC psychotic disorders, dyskinesias, and organ transplant rejection.
CC Kinase inhibitors are useful for treating diseases and disorders
CC described above.

XX
SQ Sequence 1244 AA;

Query Match 88.1%; Score 37; DB 22; Length 1244;

Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
111111
Db 651 qppplp 656

RESULT 42

AA00947
ID AA00947 standard; Protein; 1343 AA.

XX
AC AA00947;

DT 01-OCT-2001 (first entry)

DE Human bone marrow protein, SEQ ID NO: 423.

XX
KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; hemophilia.

XX
OS Homo sapiens.

XX
PN WO200153453-A2.

XX
PD 26-JUL-2001.

XX
PP 23-DEC-2000; 2000WO-US34960.

XX
PR 21-JAN-2000; 2000US-0488725.

XX
PR 25-APR-2000; 2000US-0552317.

XX
PR 09-JUL-2000; 2000US-0598042.

XX
PR 19-JUL-2000; 2000US-0620312.

XX
PR 03-AUG-2000; 2000US-0653450.

XX
PR 14-SEP-2000; 2000US-0662191.

XX
PR 19-OCT-2000; 2000US-0693036.

XX
PR 30-NOV-2000; 2000US-0250583.

XX
PA (HYSE-) HYSEQ INC.

XX
PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y,

XX
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX
PI Zhao QA, Zhou P, Drmanac RT;

XX
DR WPI; 2001-488707/53.

XX
DR N-PSDB; AA090066.

XX
PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders -
XX
PS Claim 10; Page 511-514; 648pp; English.

XX
SQ Sequence 1343 AA;

Query Match 88.1%; Score 37; DB 22; Length 1343;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 QPPRP 7
111111
Db 306 qpprip 311

RESULT 43

AAB41785

ID AAB41785 standard; Protein: 1359 AA.

AC AAB41785;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF1549 polypeptide sequence SEQ ID NO:3098.

Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
vulnery; antipariatic; antiparkinsonian; noctropic; neuroprotective;
anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
hypotensive; dermatological; immunosuppressive; antineoplastic;
antiviral; antibacterial; antifungal; antineumatic; antihyroid;
antianemic; gene therapy; cancer; proliferative disorder; hypertension;
neurodegenerative disorder; osteoarthritis; graft vs host disease;
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
bone damage; cartilage damage; antiinflammatory disease; coagulation;
thrombosis; contraceptive.

OS Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CDBA-) CDBA CORP.

XX Shinketsu RA, Leach M;

XX WPI: 2000-602362/57.

XX N-PSDB; AAC75994.

XX Novel nucleic acids and peptides derived from open reading frame X,
useful for treating e.g. cancers, proliferative disorders,
neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 2314; 5507pp; English.

AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
which represent the human ORFX open reading frames 1 to 3161. The ORFX
sequences have activities such as: cytostatic; hepatotropic; vulnery;
antipariatic; antiparkinsonian; noctropic; neuroprotective;
osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
antidiabetic; hypotensive; dermatological; immunosuppressive;
antineoplastic; antibacterial; antifungal; antineumatic;
antihyroid; and antianemic. The sequences can be used for determining
the presence of or predisposition to, or preventing or treating
pathological conditions associated with an ORFX-associated disorder. The
nucleic acids can be used to express ORFX proteins in gene therapy
vectors. The proteins and nucleic acids may be used to treat cancers,

proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hyperextension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

SQ Sequence 1359 AA;

Query Match 88.1%; Score 37; DB 21; Length 1359;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPRP 7
111111
Db 306 qpprip 311

RESULT 44

ABB71451

ID ABB71451 standard; Protein: 1475 AA.

AC ABB71451;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 41145.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmacological.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

XX N-PSDB; ABL15554.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -

XX Disclosure; SEQ ID NO 41145; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
sequences (ABL01840-ABL16175) and the encoded proteins
(ABB57737-ABB72072).
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1475 AA;

Query Match 88.1%; Score 37; DB 22; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPEPLP 7
 |||||
 Db 906 qpplp 911

RESULT 45
 ABB11890
 ID ABB11890 standard; peptide: 1554 AA.
 AC ABB11890;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human procadherin Flamingo 1 homologue, SEQ ID NO:2260.
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; actvlin;
 KW inhiblin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antistatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytosolic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnerary; antulcer.
 XX
 OS Homo sapiens.
 PN WO200157188-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US03800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 DR WPI: 2001-457740/49.
 DR N-PSDB; ABA09134.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 PS Claim 20; Page 271-272; 1963pp; English.
 XX
 XX Sequences ABB10981-ABB1230 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; actvlin- or inhiblin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be

CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.

QY 2 QPEPLP 7
 |||||
 Db 1271 qpplp 1276

Query Match 88.1%; Score 37; DB 22; Length 1554;
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 46
 ABB63122
 ID ABB63122 standard; Protein: 1629 AA.
 AC ABB63122;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 16158.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EM;
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL07225.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 16158; 21pp + Sequence listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is

FT Domain 1438..1469
 FT /note= "BGF domain"
 FT Domain 1478..1512
 FT /note= "EGF domain"
 FT Domain 1536..1553
 FT /note= "EGF domain"
 FT Domain 1542..1704
 FT /note= "Laminin G domain"
 FT Domain 1725..1736
 FT /note= "EGF domain"
 FT Domain 1792..1920
 FT /note= "Laminin G domain"
 FT Domain 1931..1962
 FT /note= "BGF domain"
 FT Domain 1966..2000
 FT /note= "EGF domain"
 FT Region 2110..2137
 FT /note= "GPCR signature"
 FT Domain 2525..2545
 FT /note= "Transmembrane domain"
 FT Region 2588..2613
 FT /note= "GPCR signature"
 FT Domain 2669..2689
 FT /note= "Transmembrane domain"
 FT Domain 2712..2730
 FT /note= "Transmembrane domain"
 FT Domain 2741..2761
 FT /note= "Transmembrane domain"
 PN WO200142285-A2.
 PD 14-JUN-2001.
 PF 05-DEC-2000; 2000WO-US32990.
 PR 10-DEC-1999; 99US-0172852.
 PR 16-DEC-1999; 99US-0172354.
 PA (INCY-) INCYTE GENOMICS INC.
 PI Yue H, Tang YT, Lal P, Burford N, Azimzal Y, Patterson C;
 PI Baughn MR, Lu DM, Shah P, Au-Young J;
 DR WPI: 2001-381632/40.
 DR N-PSDB: AAD08065.
 XX New human extracellular matrix and cell adhesion molecules and
 PT polynucleotide sequences encoding them, useful for diagnosis,
 PT prevention, treatment of genetic, autoimmune and cell proliferative
 PT disorders
 PS Claim 1; Page 113-119; 135pp; English.
 XX
 CC The present sequence is a human extracellular matrix and cell
 CC adhesion molecule (XMAP). The XMAP is used for screening a compound for
 CC effectiveness as an agonist or antagonist of XMAP. The identified agonist
 CC or antagonist are used for treating a disease or condition associated
 CC with decreased or increased expression of functional XMAP. The
 CC polynucleotides encoding XMAP are useful in somatic or germ-line gene
 CC therapy to correct a genetic deficiency, to express a conditionally
 CC lethal gene product and to express a protein which affords protection
 CC against intracellular parasites and also for diagnosis of disorders
 CC associated with expression of XMAP. They are also used for generating
 CC hybridisation probes useful in mapping the naturally occurring genomic
 CC sequences and to create knock in humanised animals (pigs) or transgenic
 CC animals (mice or rats) to model human diseases. Oligonucleotide or longer
 CC fragments derived from the polynucleotide sequences may be used as
 CC elements on a microarray. Antibodies which specifically bind XMAP may be
 CC used for the diagnosis of disorders associated with the expression of
 CC XMAP, or in assays to monitor patients being treated with XMAP. Diseases
 CC diagnosed, prevented or treated include genetic disorders such as
 CC adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's
 CC disease, myotonic dystrophy, sickle cell anaemia, thalassemia,

CC autoimmune/inflammatory disorders such as acquired immune deficiency
 CC syndrome (AIDS), Addison's disease, allergies, anaemia, asthma,
 CC atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis,
 CC glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,
 CC osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
 CC bacterial, fungal, parasitic, protozoal and helminthic infections and
 CC cell proliferative disorders such as actinic keratosis, arteriosclerosis
 CC and cancer including breast, bladder, bone marrow, brain and uterus
 CC cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.
 XX
 SO Sequence 3298 AA:
 OY 2 GPPPLP 7
 DB 3154 GPPPLP 3159
 Query Match 88.1%; Score 37; DB 22; Length 3298;
 Best Local Similarity 100.0%; Pred. No. 4.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 49
 AAE03049
 ID AAE03049 standard; Protein; 159 AA.
 AC AAE03049;
 XX 03-AUG-2001 (first entry)
 DE Human preoptic regulatory factor-2 (hPORF-2) protein #2.
 XX Human: preoptic regulatory factor-2; PORF-2; vasotropic; noctropic;
 KW neuroprotective; hormonal; anabolic; gene therapy; neurodegeneration;
 KW impotence; infertility; libido; narcolepsy; apnea; anorexia nervosa;
 KW bulimia; sleep disorder; nervous system damage.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT Protein /label= Signal_peptide
 FT /label= Mature_hPORF_2_protein
 PN WO200142464-A2.
 PD 14-JUN-2001.
 PF 30-NOV-2000; 2000WO-US30169.
 PR 08-DEC-1999; 99US-0169686.
 PA (ELIL) LILLY & CO ELI.
 PI Li XM, Rostreck PRJ, Su EW, Wang H, Wei J;
 DR N-PSDB: AAD06910.
 XX WPI: 2001-381685/40.
 PT reduced libido, narcolepsy, apnea, anorexia nervosa and bulimia
 PT
 PS Claim 9; Page 78-79; 79pp; English.
 XX
 CC The present sequence is human preoptic regulatory factor-2 (hPORF-2).
 CC hPORF2 is a neuropeptide which is used for promoting neurite outgrowth,
 CC inducing neural regeneration, promoting or inhibiting primary or
 CC secondary sexual development and altering behavioural patterns
 CC such as sleep and eating disorders. Pharmaceutical compositions
 CC comprising PORF-2 are useful for the manufacture of a medicament for the
 CC treatment or prevention of a disease in which aberrant levels of

CC endogenous PORF-2 polypeptide are partially responsible for inducing the
CC cellular effects that lead to the disease, where the disease is selected
CC from peripheral nervous system nerve damage, central nervous system
CC nerve damage, neurodegeneration, abnormal primary or secondary sexual
CC development, impotence, infertility, reduced libido, narcolepsy, apnea,
CC anorexia nervosa, and bulimia. hPORF-2 cDNA is also useful in gene
CC therapy.
XX
SQ Sequence 159 AA:

Query Match 85.7%; Score 36; DB 22; Length 159;
Best Local Similarity 71.4%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MQPPLP 7
Db 139 lepplp 145
:::|||||

RESULT 50
AAG34163
ID AAG34163 standard; Protein: 171 AA.
XX
AC AAG34163;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 41526.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
XX Zea mays subsp. mays.
OS
PM EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145224.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.7%; Score 36; DB 21; Length 171;
Best Local Similarity 71.4%; Pred. No. 4.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOPPLP 7
DB 15 lepplp 21

Search completed: August 8, 2002, 07:47:57
Job time: 130 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2002, 07:46:27 ; Search time 14.74 Seconds
(Without alignments)
45.633 Million cell updates/sec

Title: us-09-641-801-1
Perfect score: 42
Sequence: 1 MQPPPLP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :
1: PIR:1:*
2: PIR:2:*
3: PIR:3:*
4: PIR:4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	40	95.2	1776 2	G86280 protein T5E21.13 l
2	39	92.9	612 2	beta-galactosidase
3	38	90.5	536 2	DNA-(apurinic or a
4	38	90.5	542 2	protein F20B24.8 l
5	37	88.1	164 2	T15525 hypothetical prote
6	37	88.1	178 2	T20644 hypothetical prote
7	37	88.1	186 2	T22685 hypothetical prote
8	37	88.1	211 2	S55129 transcription fact
9	37	88.1	278 2	A49266 fas ligand - rat
10	37	88.1	393 2	T00575 probable WRKY-type
11	37	88.1	471 2	T21349 hypothetical prote
12	37	88.1	483 2	A25896 beta-adrenergic re
13	37	88.1	588 2	E96633 probable Serine/Th
14	37	88.1	620 2	S06733 hydroxyproline-ric
15	37	88.1	649 2	JM0809 drebrin E (clone g
16	37	88.1	651 2	T14763 hypothetical prote
17	37	88.1	731 2	JC7701 ARHGAP9 protein -
18	37	88.1	749 2	G86186 hypothetical prote
19	37	88.1	830 2	T18860 hypothetical prote
20	37	88.1	1124 2	JX0293 zinc finger protei
21	37	88.1	1125 2	T19193 E-box-binding repr
22	37	88.1	1154 2	A56242 hypothetical prote
23	37	88.1	1291 2	T17242 MEK2 protein - hu
24	37	88.1	1364 2	T00250 hypothetical prote
25	37	88.1	1360 2	T00080 FNN19 protein - ye
26	36	85.7	489 2	S23410 actin-depolymeriza
27	36	85.7	505 2	S72273 Ca2+-binding prote
28	36	85.7	542 2	JC7391 hypothetical prote
29	36	85.7	610 2	T22687

30	36	85.7	736 2	I51691 dishevelled homolo
31	36	85.7	795 2	T49835 hypothetical prote
32	36	85.7	823 2	A36378 probable transcrip
33	35	83.3	267 2	T47307 hypothetical prote
34	35	83.3	300 2	T25119 hypothetical prote
35	35	83.3	312 2	A86279 F14L17.21 protein
36	35	83.3	321 2	G86411 protein FIK23.14 l
37	35	83.3	337 2	F96507 hypothetical prote
38	35	83.3	414 2	A88485 protein F23F12.6 l
39	35	83.3	426 2	T47365 hypothetical prote
40	35	83.3	486 2	T21481 hypothetical prote
41	35	83.3	494 2	T19550 CM17 protein - mo
42	35	83.3	458 2	S52735 hypothetical prote
43	35	83.3	587 2	S49942 hypothetical prote
44	35	83.3	537 2	S72468 probable transcrip
45	35	83.3	639 2	G02919 transcription fact
46	35	83.3	666 2	T31461 probable magnesium
47	35	83.3	688 2	T09941 transcription fact
48	35	83.3	805 2	T41810 ACMPV orf66 - Bom
49	35	83.3	808 2	C72858 ACOT-66 protein -
50	35	83.3	945 2	T00746 disintegrin and me
51	35	83.3	952 2	T18900 helicase II - huma
52	35	83.3	1641 2	I38614 hypothetical prote
53	34	81.0	89 2	S72538 hypothetical prote
54	34	81.0	91 2	T45346 hypothetical prote
55	34	81.0	91 2	D70780 hypothetical prote
56	34	81.0	107 2	D71026 hypothetical prote
57	34	81.0	113 2	T17503 hypothetical prote
58	34	81.0	119 2	C40513 progesterone membr
59	34	81.0	194 2	JC5260 hypothetical prote
60	34	81.0	205 2	T35505 hypothetical prote
61	34	81.0	220 2	T10042 hypothetical prote
62	34	81.0	226 2	T22860 superoxide dismuta
63	34	81.0	227 2	A39267 hypothetical prote
64	34	81.0	289 2	T52354 hypothetical prote
65	34	81.0	335 2	T34086 hypothetical prote
66	34	81.0	360 2	T45956 hypothetical prote
67	34	81.0	364 1	T18972 phosphoprotein pho
68	34	81.0	409 2	S70704 carbon catabolite
69	34	81.0	424 2	A54964 probable glycoprot
70	34	81.0	472 2	T03169 hypothetical prote
71	34	81.0	472 2	T04855 hypothetical prote
72	34	81.0	472 2	C70853 hypothetical prote
73	34	81.0	545 2	T14632 hypothetical prote
74	34	81.0	587 2	D84426 hypothetical prote
75	34	81.0	718 1	TNBER6 77k alpha trans-in
76	34	81.0	864 2	H85335 hypothetical prote
77	34	81.0	864 2	T04518 hypothetical prote
78	34	81.0	872 2	S33015 hypothetical prote
79	34	81.0	906 2	A71438 probable resistanc
80	34	81.0	919 2	T32541 unc-5 protein - Ca
81	34	81.0	947 1	B44294 hypothetical prote
82	34	81.0	992 2	A31666 hypothetical prote
83	34	81.0	1006 2	T00050 hypothetical prote
84	34	81.0	1091 2	T34247 hypothetical prote
85	34	81.0	1107 2	T34246 membrane protein p
86	34	81.0	1182 2	T13952 probable calcium-b
87	34	81.0	1381 1	S45781 guanine nucleotide
88	34	81.0	1383 2	T13052 phosphoinositide 3
89	34	81.0	1634 1	JC5500 myoblast city prot
90	34	81.0	1970 2	T03284 variant-specific s
91	34	81.0	2228 2	T14029 hypothetical prote
92	34	81.0	2561 2	T24864 hypothetical prote
93	33	78.6	154 2	E72624 hypothetical prote
94	33	78.6	194 2	C72517 hypothetical prote
95	33	78.6	239 2	T36995 hypothetical prote
96	33	78.6	257 2	F72680 hypothetical prote
97	33	78.6	257 2	S32101 PHLPSA protein - c
98	33	78.6	322 2	B86191 hypothetical prote
99	33	78.6	322 2	S09779 hypothetical prote
100	33	78.6	324 2	S58061 amelin 2 - rat
101	33	78.6	329 2	T04244 hypothetical prote
102	33	78.6	332 2	H84443 homeodomain transc

103	33	78.6	338	2	C75353	ribosomal large su
104	33	78.6	347	2	T20472	hypothetical prote
105	33	78.6	379	2	T05441	proline-rich prote
106	33	78.6	379	2	D85257	extensin-like prote
107	33	78.6	389	2	S27200	proline-rich prote
108	33	78.6	418	2	H72203	hypothetical prote
109	33	78.6	447	2	S44809	F4489.8 protein -
110	33	78.6	462	2	E70955	hypothetical prote
111	33	78.6	474	1	ERAD41	early E2A DNA-bind
112	33	78.6	475	2	D86209	protein F2265.18 (
113	33	78.6	521	1	FOLJCA	99g polyprotein -
114	33	78.6	540	2	B96631	probable polygalac
115	33	78.6	567	2	T44928	L-ascorbate oxidas
116	33	78.6	589	2	T05074	hypothetical prote
117	33	78.6	624	2	D96494	protein F7F22.10 (
118	33	78.6	646	2	T34532	hypothetical prote
119	33	78.6	651	2	T42644	hypothetical prote
120	33	78.6	679	2	S37842	cholesterol O-acetyl
121	33	78.6	748	2	A60202	ubiquitin-protein
122	33	78.6	820	2	T46412	hypothetical prote
123	33	78.6	879	2	AC2347	hypothetical prote
124	33	78.6	1013	2	T33470	C3g protein, long
125	33	78.6	1086	2	JC7736	hypothetical prote
126	33	78.6	1250	2	T00454	hypothetical prote
127	33	78.6	1305	2	T23314	rig-1 protein - mo
128	33	78.6	1344	2	T14316	Ich1 protein - ink
129	33	78.6	1353	2	T00249	G-protein signalin
130	33	78.6	1387	2	JC5502	MEGF8 protein - hu
131	33	78.6	1737	2	T00209	tuberculous sclerosi
132	33	78.6	1784	2	A49420	prolylendopeptidas
133	33	78.6	18	2	PC2280	protein F2401.6 (1
134	33	76.2	70	2	A96650	hypothetical prote
135	33	76.2	74	2	E36800	hypothetical prote
136	33	76.2	89	2	E37731	hypothetical prote
137	33	76.2	100	2	T17126	hypothetical prote
138	33	76.2	101	2	T4506	hypothetical prote
139	33	76.2	107	2	S20693	hypothetical prote
140	33	76.2	113	2	T06155	gene dix-7 protein
141	33	76.2	114	2	E75690	hypothetical prote
142	33	76.2	117	2	E72536	hypothetical prote
143	33	76.2	121	2	B96647	hypothetical prote
144	33	76.2	125	2	T03283	synapsin II - mous
145	33	76.2	126	2	E61260	hypothetical prote
146	33	76.2	134	2	D84672	hypothetical prote
147	33	76.2	138	2	C96734	extensin precursor
148	33	76.2	139	2	S61885	glutaredoxin-like
149	33	76.2	145	2	T48552	hypothetical prote
150	32	76.2	146	2	E72695	

ALIGNMENTS

protein T5E21.13 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: G86280
 R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, R.;
 Jensen, M.F.; Hughes, B.; Hultner, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A66141; M01D:21016719
 A:Accession: G86280
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-1776 <STO>
 A:Cross-references: GB:AE005172; NID:97527720; PIDN:AAF63169.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T5E21.13
 A:Map position: 1
 Query Match 95.2%; Score 40; DB 2; Length 1776;
 Best Local Similarity 85.7%; Pred. No. 2.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOPPP 7
 Db 1643 MOPPP 1649
 RESULT 2
 E82756
 beta-galactosidase XF0840 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C:Accession: E82756
 R:Anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; M01D:20365717
 A>Note: for a complete list of authors see reference number A59328 below
 A:Accession: E82756
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-612 <STO>
 A:Cross-references: GB:AE003923; GB:AE003849; NID:99105736; PIDN:AAF83650.1; GSPDB:GN
 A:Experimental source: strain 9a5c
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Fr
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Le
 Submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Fr
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
 Rodrigues, V.; Rosa, A.J.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savr
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
 M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0840
 C:Superfamily: beta-galactosidase bga

Query Match 92.9%; Score 39; DB 2; Length 612;
 Best Local Similarity 85.7%; Pred. No. 99;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOPPP 7
 Db 356 LOPPP 362

RESULT 3
 T02441
 DNN-(apurinic or apyrimidinic site) lyase (ARP) [imported] - Arabidopsis thaliana
 N:Alternate names: hypothetical protein T26J13.5
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 02-Feb-2001
 C:Accession: T02441; A84842
 R:Bounslay, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; I
 submitted to the EMBL Data Library, June 1998
 A:Description: Arabidopsis thaliana chromosome II BAC T26J13 genomic sequence.
 A:Reference number: Z14673
 A:Accession: T02441

A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-536 <R0D>
 A:Cross-references: EMBL:AC004625; NID:g3241939; PID:g3241944
 A:Experimental source: cultivar Columbia
 A:Authors: R.L. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, R.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; Nuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: AB4420; M0ID:20083487
 A:Accession: AB4842
 A:Molecule type: DNA
 A:Status: preliminary
 A:Residues: 1-536 <STO>
 A:Cross-references: GB:AE002093; NID:g3241944; PIDN:AA023731.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: AT2G41460; T26U13.5
 A:Map position: 2
 A:introns: 18/2; 78/3; 110/2; 137/1; 210/1; 317/3; 354/3; 400/3; 416/3; 443/3; 497/1

Query Match 90.5%; Score 38; DB 2; Length 536;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 Db 264 MRPPPLP 270

RESULT 4
 H86239
 protein F20R24.8 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: H86239
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Maritali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: AB6141; M0ID:21016719
 A:Accession: H86239
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-542 <STO>
 A:Cross-references: GB:AE005172; NID:g6573750; PIDN:AAFI670.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F20R24.8
 A:Map position: 1

Query Match 90.5%; Score 38; DB 2; Length 542;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 Db 105 VOPPLP 111

RESULT 5
 T15525
 hypothetical protein C16B8.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
 C:Accession: T15525

R:Bentley, D.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans consold C16B8.
 A:Reference number: Z18365
 A:Accession: T15525
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-164 <BEN>
 A:Cross-references: EMBL:U41031; NID:g1098982; PID:g1098985; PIDN:AAA82620.1; CESP:C1
 C:Genetics:
 A:Gene: CESP:C16B8.3
 C:Superfamily: proline-rich protein

Query Match 88.1%; Score 37; DB 2; Length 164;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 Db 97 QPPPLP 102

RESULT 6
 T00644
 hypothetical protein F316.7 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 22-Oct-1999
 C:Accession: T00644
 R:Federspiel, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Arau
 R; Vysotskaia, V.S.; Yu, G.; Ecker, J.; Theologis, A.; Davis, R.W.
 submitted to the EMBL Data Library, February 1998
 A:Reference number: Z14197
 A:Accession: T00644
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-178 <FED>
 A:Cross-references: EMBL:AC002396; NID:g2749918; PID:g2829866; GSPDB:GN00059; ATSP:F3
 C:Genetics:
 A:Gene: ATSP:F316.7
 A:Map position: 1
 A:introns: 50/3; 87/1
 C:Superfamily: Arabidopsis thaliana hypothetical protein F316.7

Query Match 88.1%; Score 37; DB 2; Length 178;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 Db 162 QPPPLP 167

RESULT 7
 T22685
 hypothetical protein F55A11.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T22685
 R:Kershaw, J.
 submitted to the EMBL Data Library, May 1996
 A:Reference number: Z19600
 A:Accession: T22685
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-186 <WIL>
 A:Cross-references: EMBL:472511; PIDN:CAA96655.1; GSPDB:GN00023; CESP:F55A11.1
 A:Experimental source: clone F55A11
 C:Genetics:
 A:Gene: CESP:F55A11.1
 A:Map position: 5
 A:introns: 33/3; 59/2; 146/2

Query Match 88.1%; Score 37; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
|||||
Db 127 QPPPLP 132

RESULT 8

S55129
transcription factor RGM1 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YM8010.12c; protein YMR182c
C:Species: Saccharomyces cerevisiae
C:Date: 08-Jul-1995 #sequence_revision-01-Sep-1995 #text_change 21-Jul-2000
C:Accession: S55129; S17249
R:Churcher, C.M.
submitted to the EMBL Data Library, June 1995
A:Accession: S55129
A:Molecule type: DNA
A:Residues: 1-211 <CHD>
A:Cross-references: EMBL:Z49808; NID:g854440; PID:g854452; MIPS:YMR182c
A:Experimental source: strain AB972
R:Estreich, F.
Nucleic Acids Res. 19, 4873-4877, 1991
A:Title: The yeast putative transcriptional repressor RGM1 is a proline-rich zinc finger
A:Reference number: S17249; MUID:92020118
A:Accession: S17249
A:Molecule type: DNA
A:Residues: 1-113, 'V', 115-211 <EST>
A:Cross-references: EMBL:X59861; NID:g4317; PIDN:CAAA2521.1; PID:g4318
C:Genetics:
A:Gene: SGD:RGM1
A:Cross-references: SGD:S0004794; MIPS:YMR182c
A:Map position: 13R
C:Keywords: DNA binding; transcription factor; zinc finger

Query Match 88.1%; Score 37; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
|||||
Db 154 QPPPLP 159

RESULT 9

A49266
fas ligand - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C:Accession: A49266
R:Stude, T.; Takahashi, T.; Golstein, P.; Nagata, S.
Cell 75, 1169-1178, 1993
A:Title: Molecular cloning and expression of the Fas ligand, a novel member of the tumor
A:Reference number: A49266; MUID:94084792
A:Accession: A49266
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-278 <SND>
A:Cross-references: GB:U03470; NID:g440178; PIDN:AACS2129.1; PID:g440179
C:Keywords: glycoprotein; transmembrane protein

Query Match 88.1%; Score 37; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7

Db 60 QPPPLP 65
|||||

RESULT 10

T00575
probable WRKY-type DNA binding protein [imported] - Arabidopsis thaliana
N:Alternate names: probable DNA-binding protein T27E13.1
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001
C:Accession: T00575; B84706
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K
submitted to the EMBL Data Library, May 1998
A:Description: Arabidopsis thaliana chromosome II BAC T27E13 genomic sequence.
A:Reference number: Z14178
A:Accession: T00575
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-393 <ROD>
A:Cross-references: EMBL:AC004165; NID:g3150396; PID:g3150397
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, L.; Tallon,
euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: B84706
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <STO>
A:Cross-references: GB:AE002093; NID:g2347191; PIDN:AC16930.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g30250; T27E13.1
A:Map position: 2
A:Introns: 79/3; 117/3; 294/2; 348/2

Query Match 88.1%; Score 37; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 136+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
|||||
Db 50 QPPPLP 55

RESULT 11

T21349
hypothetical protein F25H2.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T21349
R:Wilkinson, J.
submitted to the EMBL Data Library, September 1996
A:Reference number: Z19411
A:Accession: T21349
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-471 <III>
A:Cross-references: EMBL:Z79754; PIDN:CA802096.1; GSPDB:GN00019; CESP:F25H2.8
A:Experimental source: clone F25H2
C:Genetics:
A:Gene: CESP:F25H2.8
A:Map position: 1
A:Introns: 26/3; 138/3; 302/3; 353/3; 379/2
C:Superfamily: Caenorhabditis elegans hypothetical protein F25H2.8

Query Match 88.1%; Score 37; DB 2; Length 471;
Best Local Similarity 71.4%; Pred. No. 156+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
DB 451 LOPPLP 457

RESULT 12

A:Accession: A25896
A:Residues: 1-483 <VAR>
A:Molecule type: mRNA
A:Cross-references: GB:M14379; NID:g213891; PIDD:AAA9627.1; PID:g213892
A:Experimental source: erythrocyte
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 88.1%; Score 37; DB 2; Length 483;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPEPLP 7
DB 256 QPEPLP 261

RESULT 13

E96633
Probable Serine/Threonine protein kinase F8A5.31 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E96633
R:Thelodopsis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chinn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, T.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: E96633
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-588 <STO>
A:Cross-references: GB:AE05173; NID:g2462749; PIDD:AA871968.1; GSPDB:GN00141
C:Genetics:
A:Gene: F8A5.31
A:Map position: 1

Query Match 88.1%; Score 37; DB 2; Length 588;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPEPLP 7
DB 550 QPEPLP 555

RESULT 14

S06733
hydroxyproline-rich glycoprotein precursor - common tobacco
C:Species: Nicotiana tabacum (common tobacco)

C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jul-2000
C:Accession: S06733
R:Keller, B.; Lamb, C.J.
Genes Dev. 3, 1639-1646, 1989
A:Title: Specific expression of a novel cell wall hydroxyproline-rich glycoprotein gene
A:Reference number: S06733; MUID:90128263
A:Accession: S06733
A:Molecule type: DNA
A:Residues: 1-620 <KE>
A:Cross-references: EMBL:X13885; NID:g19866; PIDD:CAA32090.1; PID:g19867
C:Superfamily: hydroxyproline-rich glycoprotein
C:Keywords: glycoprotein

Query Match 88.1%; Score 37; DB 2; Length 620;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPEPLP 7
DB 425 QPEPLP 430

RESULT 15

JN0809
drebrin E (clone gdh13) - human
C:Species: Homo sapiens (man)
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000
C:Accession: JN0809
R: Toda, M.; Shitao, T.; Minoshima, S.; Shimizu, N.; Toyo, S.; Uyemura, K.
Biochem. Biophys. Res. Commun. 196, 468-472, 1993
A:Title: Molecular cloning of cDNA encoding human drebrin E and chromosomal mapping
A:Reference number: JN0809; MUID:94030036
A:Accession: JN0809
A:Molecule type: mRNA
A:Residues: 1-649 <TO>
A:Cross-references: DDBJ:D17530; NID:g498650; PIDD:BA04480.1; PID:g498651
A:Experimental source: brain
C:Comment: This protein is an actin-binding protein.
C:Genetics:
A:Map position: 5
C:Keywords: actin binding

Query Match 88.1%; Score 37; DB 2; Length 649;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPEPLP 7
DB 363 QPEPLP 368

RESULT 16

T14763
hypothetical protein DKFZp434D064.1 - human
C:Species: Homo sapiens (man)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14763
R: Mambuti, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A:Reference number: Z18181
A:Accession: T14763
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-651 <RAM>
A:Cross-references: EMBL:AL110225
A:Experimental source: adult testis; clone DKFZp434D064
C:Genetics:
A:Note: DKFZp434D064.1

Query Match 88.1%; Score 37; DB 2; Length 651;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 QPPPLP 7
|||||
Db 365 QPPPLP 370

RESULT 17
JC7701
ARHGAP9 protein - human
C:Species: Homo sapiens (man)
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: JC7701
R:Furukawa, Y.; Kawasoe, T.; Dalgo, Y.; Nishiwaki, T.; Ishiguro, H.; Takahashi, M.; Kite
B:Biochem. Biophys. Res. Commun. 284, 643-649, 2001
A:Title: Isolation of a novel human gene, ARHGAP9, encoding a Rho-GTPase activating pro
A:Reference number: JC7701; MUID:21290655; PMID:11396949
A:Accession: JC7701
A:Molecule type: mRNA
A:Residues: 1-731 <FUK>
A:Cross-references: DDBJ:AB051853
C:Comment: This protein, a novel member of the Rho-GTPase activating protein (Rho-GAP) f
form of Cdc42hs, Rac1, and RhoA in vitro, and it is also involved in adhesion of hemato
C:Genetics:
A:Gene: arhgap9

Query Match 88.1%; Score 37; DB 2; Length 731;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
|||||
Db 131 QPPPLP 136

RESULT 18
G86186
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G86186
R:Phelogeis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Hughes, B.; Huizar, L.
ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mailli, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: G86186
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-749 <STO>
A:Cross-references: GB:AE005172; NID:92388574; PIDN:AB071455.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 88.1%; Score 37; DB 2; Length 749;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MQPPLP 7
|||||
Db 221 MSPPLP 227

RESULT 19

T18860
hypothetical protein C02C6.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T18860
R:Swindurne, J.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19032
A:Accession: T18860
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-830 <MIL>
A:Cross-references: EMBL:Z79596; PIDN:CAB01857.1; GSPDB:GN00028; CESP:C02C6.1
A:Experimental source: clone C02C6
C:Genetics:
A:Gene: CESP:C02C6.1
A:Map position: X
A:Introns: 56/2; 131/1; 333/2; 501/2; 593/2; 686/3; 815/2
C:Superfamily: human dynamin II; pleckstrin repeat homology

Query Match 88.1%; Score 37; DB 2; Length 830;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
|||||
Db 747 QPPPLP 752

RESULT 20
JX0293
zinc finger protein ARRB6 - human
C:Species: Homo sapiens (man)
C:Date: 03-Feb-1994 #sequence_revision 20-Aug-1994 #text_change 28-May-1999
C:Accession: JX0293
R:Watanabe, Y.; Kawakami, K.; Hirayama, Y.; Nagano, K.
J. Biochem. 114, 849-855, 1993
A:Title: Transcription factors positively and negatively regulating the Na,K-ATPase
A:Reference number: JX0293; MUID:94186507
A:Accession: JX0293
A:Molecule type: mRNA
A:Residues: 1-1114 <MAT>
A:Cross-references: GB:D15050; NID:9457560; PIDN:BA03646.1; PID:dl004160; PID:95585;
C:Keywords: zinc finger
F:172-193/Region: zinc finger
F:202-222/Region: zinc finger
F:242-262/Region: zinc finger
F:270-292/Region: zinc finger
F:906-926/Region: zinc finger
F:934-954/Region: zinc finger
F:962-981/Region: zinc finger
F:989-1123/Region: glutamic acid-rich

Query Match 88.1%; Score 37; DB 2; Length 1124;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
|||||
Db 557 QPPPLP 562

RESULT 21
T19193
hypothetical protein C11F4.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T19193
R:Northmore, B.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19087

A:Accession: T19193
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1125 <WIL>
 A:Cross-references: EMBL:Z81015; PIDN:CAB02658.1; GSPDB:GN00028; CESP:C11E4.6
 A:Experimental source: clone C11E4
 C:Genetics:
 A:Gene: CESP:C11E4.6
 A:Map position: X
 A:introns: 42/2; 155/3; 232/3; 349/1; 454/1; 558/1; 601/1; 774/3; 831/3; 906/1; 949/3; 9

Query Match 88.1%; Score 37; DB 2; Length 1125;
 Best Local Similarity 85.7%; Pred. No. 3.7e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 Db 987 MSPPLP 993
 |||||

RESULT 22
 A56242
 E-box-binding repressor ZEB - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 07-May-1999
 C:Accession: A56242
 R:Genetta, T.; Ruzinskiy, D.; Kadesch, T.
 Mol. Cell. Biol. 14, 6153-6163, 1994
 A:Title: Displacement of an E-box-binding repressor by basic helix-loop-helix proteins:
 A:Reference number: A56242; MUID:94344126
 A:Accession: A56242
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1154 <GEN>
 A:Cross-references: GB:U19969
 C:Keywords: DNA binding; zinc finger

Query Match 88.1%; Score 37; DB 2; Length 1154;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPLP 7
 Db 591 OPPLP 596
 |||||

RESULT 23
 T17242
 hypothetical protein DKFZp586B1417.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T17242
 R:Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, September 1999
 A:Reference number: Z18722
 A:Accession: T17242
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1291 <KOE>
 A:Cross-references: EMBL:AL117448
 A:Experimental source: adult uterus; clone DKFZp586B1417
 C:Genetics:
 A:Note: DKFZp586B1417.1

Query Match 88.1%; Score 37; DB 2; Length 1291;
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPLP 7
 |||||

DB 238 QPPLP 243
 |||||

RESULT 24
 T00250
 MEGF2 protein - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
 C:Accession: T00250
 R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
 Genomics 51, 27-34, 1998
 A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
 A:Reference number: Z14126; MUID:98360089
 A:Accession: T00250
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1364 <NAK>
 A:Cross-references: EMBL:AB011536; NID:G3449297; PIDN:BAA32464.1; PID:G3449298
 A:Experimental source: brain; clone HG1044
 C:Genetics:
 A:Gene: MEGF2
 A:Map position: 3p21.2-p24.1
 C:Superfamily: unassigned EGF-related proteins; EGF homology; laminin-type EGF-like h
 F:1-28/Domain: EGF homology (fragment) <EGF>
 F:32-66/Domain: EGF homology <EGF>
 F:124-169/Domain: laminin-type EGF-like homology <LEG>

Query Match 88.1%; Score 37; DB 2; Length 1364;
 Best Local Similarity 100.0%; Pred. No. 4.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPLP 7
 Db 1220 QPPLP 1225
 |||||

RESULT 25
 T00080
 hypothetical protein KIA0522 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
 C:Accession: T00080
 R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara,
 DNA Res. 5, 31-39, 1998
 A:Title: Prediction of the coding sequences of unidentified human genes. IX. The com
 A:Reference number: Z14086; MUID:98290545
 A:Accession: T00080
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1560 <NAG>
 A:Cross-references: EMBL:AB011094; NID:G3043567; PIDN:BAA25448.1; PID:G3043568
 A:Experimental source: brain; clone HG1393
 C:Genetics:
 A:Note: KIA0522

Query Match 88.1%; Score 37; DB 2; Length 1560;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPLP 7
 Db 1363 QPPLP 1368
 |||||

RESULT 26
 S23410
 FUM19 protein - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YAL034c
 C:Species: Saccharomyces cerevisiae
 C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 29-Oct-1999
 C:Accession: S23410; S1986; S19060

R.Harris, S.D.; Cheng, J.; Pugh, T.A.; Pringle, J.R.
 J. Mol. Biol. 225, 53-65, 1992
 A:Title: Molecular analysis of Saccharomyces cerevisiae chromosome I. On the number of
 A:Reference number: S23409; MUID:92260538
 A:Accession: S23410
 A:Molecule type: DNA
 A:Residues: 1-489 <HAR>
 A:Cross-references: EMBL:X62577; NID:95240; PIDN:CA44456.1; PID:95242
 R.Bussey, H.; Kaback, D.B.; Zhong, W.; Vo, D.T.; Clark, M.W.; Fortin, N.; Hall, J.; Ouel
 submitted to the EMBL Data Library, August 1994
 A:Description: The sequence of chromosome 1 of Saccharomyces cerevisiae.
 A:Reference number: S51956
 A:Accession: S51986
 A:Molecule type: DNA
 A:Residues: 1-489 <BUS>
 A:Cross-references: EMBL:U12980; MIPS:YAL034C
 C:Genetics:
 A:Gene: SGD:FUN19
 A:Cross-references: SGD:S0002134; MIPS:YAL034C
 A:Map position: 1L

Query Match 85.7%; Score 36; DB 2; Length 489;
 Best Local Similarity 85.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 1 |||||
 Db 243 MPPPLP 249

RESULT 27
 S72273
 actin-depolymerizing protein N-WASP, brain - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 23-Apr-1998 #sequence_revision 01-May-1998 #text_change 05-Nov-1999
 C:Accession: S72273
 R.Miki, H.; Miura, K.; Takenawa, T.
 EMBO J. 15, 5326-5335, 1996
 A:Title: N-WASP, a novel actin-depolymerizing protein, regulates the cortical cytoskeleton
 A:Reference number: S72273; MUID:97050838
 A:Accession: S72273
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-505 <MIK>
 A:Cross-references: EMBL:D67066; NID:91644231; PIDN:BA11082.1; PID:d1011742; PID:916442
 A:Experimental source: brain

Query Match 85.7%; Score 36; DB 2; Length 505;
 Best Local Similarity 85.7%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 1 |||||
 Db 342 MPPPLP 348

RESULT 28
 JC7391
 Ca2+-binding protein, DD4 - prawn
 N:Alternate names: DD4 protein
 C:Species: Marsupeneus japonicus
 C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 18-Jul-2001
 C:Accession: JC7391
 R.Endo, H.; Persson, P.; Watanabe, T.
 Biochem. Biophys. Res. Commun. 276, 286-291, 2000
 A:Title: Molecular cloning of the crustacean DD4 cDNA encoding a Ca2+ -binding protein.
 A:Reference number: JC7391; MUID:20462952; PMID:11006119
 A:Accession: JC7391
 A:Molecule type: mRNA
 A:Residues: 1-542 <END>
 C:Comment: This protein, acidic and proline-rich, is involved in the calcification process

C:Genetics:
 A:Gene: dd4
 C:Keywords: calcium binding

Query Match 85.7%; Score 36; DB 2; Length 542;
 Best Local Similarity 71.4%; Pred. No. 2.5e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 1 |||||
 Db 357 IQPPPLP 363

RESULT 29
 T22687
 hypothetical protein F55A11.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
 C:Accession: T22687
 R.Kershaw, J.
 submitted to the EMBL Data Library, May 1996
 A:Reference number: Z19600
 A:Accession: T22687
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-610 <MIK>
 A:Cross-references: EMBL:Z72511; PIDN:CA96657.1; GSPDB:GN00023; CESP:F55A11.3
 A:Experimental source: Clone F55A11
 C:Genetics:
 A:Gene: CESP:F55A11.3
 A:Map position: 5
 A:Introns: 76/3; 221/1; 275/2; 338/1; 424/3
 C:Superfamily: RING finger homology
 F:288-338/Domain: RING finger homology <RRR>

Query Match 85.7%; Score 36; DB 2; Length 610;
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 1 |||||
 Db 444 MPPPLP 450

RESULT 30
 I51691
 dishevelled homolog - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C:Accession: I51691
 R.Sokol, S.Y.; Ringgenmuth, J.; Perrimon, N.; Itoh, K.
 Development 121, 1637-1647, 1995
 A:Title: Dorsalizing and neuralizing properties of xds, a maternally expressed Xenopus
 A:Reference number: I51691; MUID:95324391
 A:Accession: I51691
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-736 <SOK>
 A:Cross-references: EMBL:U01552; NID:9945109; PIDN:AA00688.1; PID:9945110
 C:Genetics:
 A:Gene: xds
 C:Superfamily: GLGF domain homology
 F:260-336/Domain: GLGF domain homology <GLG4>

Query Match 85.7%; Score 36; DB 2; Length 736;
 Best Local Similarity 85.7%; Pred. No. 3.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 1 |||||

Db 679 MPPPLP 685

RESULT 31

T49835
Hypothetical protein B24H17.210 [Imported] - Neurospora crassa

C:Species: Neurospora crassa

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C:Accession: T49835

R:Schulte, U.; Allyn, V.; Hohenseel, J.; Brandt, P.; Farthmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000

A:Reference number: 225022

A:Accession: T49835

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-795 <SCH>

A:Cross-references: EMBL:AL356815, GSPDB:GN00116, NCSP:B24H17.210

A:Experimental source: BAC clone B24H17; strain OR74A

C:Genetics:

A:Gene: NCSP:B24H17.210

A:Map position: 6

Query Match 85.7%; Score 36; DB 2; Length 795;

Best Local Similarity 85.7%; Pred. No. 3.6e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MPPPLP 7
:|||||

Db 419 MPPPLP 425

RESULT 32

A36378
Probable transcription activator nuc-1 - Neurospora crassa

C:Species: Neurospora crassa

C:Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 21-Jul-2000

C:Accession: A36378

R:Kang, S.; Metzberg, R.L.

Mol. Cell. Biol. 10, 5839-5848, 1990

A:Title: Molecular analysis of nuc-1(+), a gene controlling phosphorus acquisition in Ne

A:Reference number: A36378; MUID:91042513

A:Accession: A36378

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-823 <KAN>

A:Cross-references: GB:M37700; NID:g168851; PIDN:AAA33603.1; PID:g168852

C:Keywords: DNA binding; nucleus; transcription regulation

Query Match 85.7%; Score 36; DB 2; Length 823;

Best Local Similarity 85.7%; Pred. No. 3.8e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MPPPLP 7
:|||||

Db 433 MPPPLP 439

RESULT 33

T47307
Hypothetical protein T32A11.60 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T47307

R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.

submitted to the Protein Sequence Database, April 2000

A:Reference number: 224459

A:Accession: T47307

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-267 <RIE>

A:Cross-references: EMBL:AL138653

A:Experimental source: cultivar Columbia; BAC clone T32A11

C:Genetics:

A:Map position: 3

A:Introns: 53/3

A:Note: T32A11.60

Query Match 83.3%; Score 35; DB 2; Length 267;

Best Local Similarity 71.4%; Pred. No. 1.7e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MPPPLP 7
:|||||

Db 52 LQPPPP 58

RESULT 34

T25119
Hypothetical protein T22C8.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T25119

R:Thomas, K.

submitted to the EMBL Data Library, April 1995

A:Reference number: 219983

A:Accession: T25119

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-300 <WIL>

A:Cross-references: EMBL:Z49071; PIDN:CA88878.1; GSPDB:GN00020; CESP:T22C8.6

A:Experimental source: clone T22C8

C:Genetics:

A:Gene: CESP:T22C8.6

A:Map position: 2

A:Introns: 38/1; 69/3; 137/3; 185/2; 206/3

Query Match 83.3%; Score 35; DB 2; Length 300;

Best Local Similarity 71.4%; Pred. No. 1.9e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 MPPPLP 7
:|||||

Db 39 IEPPPLP 45

RESULT 35

A86279
F14L17.21 protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001

C:Accession: A86279

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alor

Chin, C.W.; Chung, M.K.; Conn, L.; Conway A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luco, J.S.; Malt, R.; Maritz

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallc

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: A86279

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-312 <STO>

A:Cross-references: GB:AE005172; NID:g7262686; PIDN:AAF3944.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 83.3%; Score 35; DB 2; Length 312;

Best Local Similarity 83.3%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0;
QY 2 MOPPLP 7
:|||||
Db 51 QPPMP 56

RESULT 36
G86411
protein FIK23.14 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G86411
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewart, R.;
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; M0ID:21016719
A:Accession: G86411
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-321 <STO>
A:Cross-references: GB:AE005172; NID:q10764855; PIDN:AAF24546.2; GSPDB:GN00141
C:Genetics: FIK23.14
A:Gene: FIK23.14
A:Map position: 1

Query Match 83.3%; Score 35; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 2.1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 1 MOPPLP 6
:|||||
Db 152 MOPPLP 157

RESULT 37
F96507
hypothetical protein T12C22.12 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F96507
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, R.;
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; M0ID:21016719
A:Accession: F96507
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-337 <STO>
A:Cross-references: GB:AE005173; NID:q8655395; PIDN:AAF78268.1; GSPDB:GN00141
C:Genetics: T12C22.12
A:Gene: T12C22.12
A:Map position: 1

Query Match 83.3%; Score 35; DB 2; Length 337;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MOPPLP 7
:|||||
Db 52 LQPPPP 58

RESULT 38
A88485
protein F23F12.6 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: A88485
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; M0ID:99069613; PMID:9851916
A>Note: see webites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_e
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: A88485
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-414 <STO>
A:Cross-references: GB:chr_III; PIDN:AAA20608.1; PID:9529215; GSPDB:GN00021; CESP:F23
C:Genetics: F23F12.6
A:Gene: F23F12.6
A:Map position: 3
C:Superfamily: ATP-dependent 26S proteinase; FltSH/SEC18/CD48-type ATP-binding domain

Query Match 83.3%; Score 35; DB 2; Length 414;
Best Local Similarity 71.4%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 0;

QY 1 MOPPLP 7
:|||||
Db 20 LRPPPLP 26

RESULT 39
T47365
hypothetical protein F7M19.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47365
R:Yakutara, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, J.
Mayer, K.F.X.
Submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24458
A:Accession: T47365
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-426 <NTA>
A:Cross-references: EMBL:AL136643
A:Experimental source: cultivar Columbia; BAC clone F7M19
C:Genetics:
A:Map position: 3
A:introns: 53/3; 213/3
A>Note: F7M19.80

Query Match 83.3%; Score 35; DB 2; Length 426;
Best Local Similarity 71.4%; Pred. No. 2.7e+02; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 1;

QY 1 MOPPLP 7
:|||||
Db 52 LQPPPP 58

RESULT 40
T21481
hypothetical protein F28C6.1 - Caenorhabditis elegans

C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T21481
 R:Burton, J.
 submitted to the EMBL Data Library, December 1995
 A:Reference number: 219429
 A:Accession: T21481
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-486 <M12>
 A:Cross-references: EMBL:Z68315; PIDN:CAA92669.1; GSPDB:GN00020; CESP:F28C6.1
 A:Experimental source: clone F28C6
 C:Genetics:
 A:Gene: CESP:F28C6.1
 A:Map position: 2
 A:Introns: 86/3; 276/2; 363/2

Query Match 83.3%; Score 35; DB 2; Length 486;
 Best Local Similarity 83.3%; Pred. No. 3.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 |||||
 Db 96 QPPMP 101

RESULT 41
 T19550
 hypothetical protein D1054.15 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
 C:Accession: T19550; T20311
 R:Wilkinson, J.
 submitted to the EMBL Data Library, June 1996
 A:Reference number: 219140
 A:Accession: T19550
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-494 <M12>
 A:Cross-references: EMBL:Z73970; PIDN:CAA98247.1; GSPDB:GN00023; CESP:D1054.15
 A:Experimental source: clone C29A12
 R:Matthews, P.
 submitted to the EMBL Data Library, June 1996
 A:Reference number: 219255
 A:Accession: T20311
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-494 <M12>
 A:Cross-references: EMBL:Z74030; PIDN:CAA98448.1; GSPDB:GN00023; CESP:D1054.15
 C:Genetics:
 A:Gene: CESP:D1054.15
 A:Map position: 5
 A:Introns: 22/1; 61/2; 249/3; 475/3
 C:Superfamily: PRL1 protein; WD repeat homology

Query Match 83.3%; Score 35; DB 2; Length 494;
 Best Local Similarity 83.3%; Pred. No. 3.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 |||||
 Db 53 QPPMP 58

RESULT 42
 S52735
 CML1R protein - mouse
 C:Species: *Mus musculus* (house mouse)
 C:Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 05-Nov-1999
 C:Accession: S52735

R.Schmitt-Wrede, H.P.; Wrehke, C.; Oino, Z.D.; Heischkamp, H.; Benten, W.P.M.; Wunde
 submitted to the EMBL Data Library, March 1995
 A:Description: Testosterone-Induced Immunosuppression of Plasmodium chabaudi malaria:
 A:Reference number: S52735
 A:Accession: S52735
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-348 <SCH>
 A:Cross-references: EMBL:X85802; NID:g1332610; PIDN:CAA59797.1; PID:g758299

Query Match 83.3%; Score 35; DB 2; Length 548;
 Best Local Similarity 85.7%; Pred. No. 3.5e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MQPPPLP 7
 |||||
 Db 425 MQPPPP 431

RESULT 43
 S49942
 hypothetical protein Y11036w - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein Y19905.12
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 13-Jan-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
 R:Odell, C.; Bowman, S.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: S49931
 A:Accession: S49942
 A:Molecule type: DNA
 A:Residues: 1-587 <ODE>
 A:Cross-references: GB:Z47047; EMBL:Z46861; NID:g603997; PID:g763310; GSPDB:GN00009;
 C:Genetics:
 A:Gene: MIPS:Y11036w
 A:Map position: 9L
 C:Superfamily: fos/jun DNA-binding domain homology
 F:420-460/Domain: fos/jun DNA-binding domain homology <F3D>

Query Match 83.3%; Score 35; DB 2; Length 587;
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 |||||
 Db 289 QPPMP 294

RESULT 44
 S72468
 Probable transcription factor grisea - *Podospira anserina*
 C:Species: *Podospira anserina*
 C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 21-Jul-2000
 C:Accession: S72468
 R:Osiewicz, H.D.; Nuber, U.
 Mol. Gen. Genet. 252, 115-124, 1996
 A:Title: GRISA, a putative copper-activated transcription factor from *Podospira anse*
 A:Reference number: S72468; MUID:96397516
 A:Accession: S72468
 A:Molecule type: DNA
 A:Residues: 1-597 <OSI>
 A:Cross-references: EMBL:X89429; NID:g1531572; PIDN:CAA61598.1; PID:g1531573
 A:Experimental source: strain S
 C:Genetics:
 A:Gene: grisea
 A:Map position: 1
 A:Introns: 11/2
 C:Function:
 A:Description: Involved in differentiation and senescence
 C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homolog
 C:Keywords: copper binding; DNA binding; transcription factor

F:173-349/Region: proline-rich
F:11/Blinding site: copper (Cys) #status predicted

Query Match 83.3%; Score 35; DB 2; Length 597;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
|||||
DB 331 QPPMP 336

RESULT 45

transcription factor ZFM1 - human

C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 01-Dec-2000

C:Accession: G02919; 154371; 168667; 168669

R:Breviario, F.
Submitted to GenBank, December 1995

A:Reference number: H01954

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-639 <BRE>
A:Cross-references: GB:L49380; NID:91405420; PIDN:AA04033.1; PID:91405421

R:Yoda, T.; Iida, A.; Miwa, T.; Nakamura, Y.; Imai, T.
Hum. Mol. Genet. 3, 465-470, 1994

A:Title: Isolation and characterization of a novel gene encoding nuclear protein at a 1c

A:Reference number: 154371; MUID:94282041

A:Accession: 154371

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-268, 'G', 270-347, 'A', 349-376, 'W', 378-586, 'NSIECICLILTLTQLPLPRGSDSPRRH

A:Cross-references: GB:D26120; NID:9785995; PIDN:BA05117.1; PID:9785997

A:Accession: 168667

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 116-138 <RE3>
A:Cross-references: GB:D26122; NID:9473832; PIDN:BA05119.1; PID:9786000

C:Genetics:
A:Gene: GDB:ZNF162; ZFM1

A:Cross-references: GDB:320035; OMIM:601516

A:Map position: 11q13-11q13

A:Introns: 528/1

Query Match 83.3%; Score 35; DB 2; Length 639;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MPPPLP 7
|||||
DB 425 MPPMP 431

RESULT 46

T31461

probable magnesium chelatase (EC 4.99.1.-) chain D bchd - Helicobacillus mobilis

C:Species: Helicobacillus mobilis

C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000

C:Accession: T31461

R:Xiong, J.; Inoue, K.; Bauer, C.E.

Proc. Natl. Acad. Sci. U.S.A. 95, 14851-14856, 1998

A:Title: Tracking molecular evolution of photosynthesis by characterization of a major P

A:Reference number: 221036; MUID:99061957

A:Accession: T31461

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-666 <XIO>
A:Cross-references: EMBL:AF080002; NID:93820536; PID:93820559; PIDN:AA04032.1
C:Genetics:
A:Gene: bchd
C:Keywords: lyase

Query Match 83.3%; Score 35; DB 2; Length 666;
Best Local Similarity 85.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MPPPLP 7
|||||
DB 334 MPPMP 340

RESULT 47

transcription factor VPI - Craterostigma plantagineum

N:Alternate names: VPI homolog
C:Species: Craterostigma plantagineum

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 18-Aug-2000

C:Accession: T09941

R:Chandler, J.W.; Bartels, D.
Mol. Gen. Genet. 256, 539-546, 1997

A:Title: Structure and function of the vpi gene homologue from the resurrection plant

A:Reference number: 216899; MUID:98075359

A:Accession: T09941

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-688 <CHAS>

A:Cross-references: EMBL:AJ000552; NID:92288898; PIDN:CA04184.1; PID:92288899

A:Experimental source: seed

C:Genetics:
A:Gene: vpi

A:Function:
A:Description: seed-specific transcriptional activator

C:Superfamily: rice transcription factor vpi

C:Keywords: DNA binding; seed; transcription factor; transcription regulation

Query Match 83.3%; Score 35; DB 2; Length 688;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
|||||
DB 339 QPPMP 344

RESULT 48

T41810

ACMPV orf66 - Bombyx mori nuclear polyhedrosis virus (isolate T3)

C:Species: Bombyx mori nuclear polyhedrosis virus, BMSNPV

A:Variety: isolate T3

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000

C:Accession: T41810

R:Gomi, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999

A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.

A:Reference number: 222020; MUID:99281911

A:Accession: T41810

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-805 <RAM>

A:Cross-references: EMBL:L33180; PIDN:AA063739.1

A:Experimental source: isolate T3

C:Genetics:
A:Note: Orf_54

Query Match 83.3%; Score 35; DB 2; Length 805;
 Best Local Similarity 85.7%; Pred. No. 5.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 MOPPLP 7
 |||||
 DB 106 MQPPPP 112

RESULT 49

C72858
 Acoorf-66 protein - Autographa californica nuclear polyhedrosis virus
 C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV
 A:Note: dsDNA virus
 C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 12-Nov-1999
 C:Accession: C72858
 R:Ayres, M.D.; Howard, S.C.; Kuzlo, J.; Lopez-Ferber, M.; Possee, R.D.
 Virology 202, 586-605, 1994
 A:title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus
 A:Reference number: A72850; MUID:94303173
 A:Accession: C72858
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-808 <AYR>
 A:Cross-References: GB:L22858; NID:9510708; PIDN:AAA66696.1; PID:9559135
 C:Genetics:
 A:Gene: Acoorf-66

Query Match 83.3%; Score 35; DB 2; Length 808;
 Best Local Similarity 85.7%; Pred. No. 5.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 |||||
 DB 104 MQPPPP 110

RESULT 50

T00746
 hypothetical protein Atg40820 [imported] - Arabidopsis thaliana
 N:Alternate names: hypothetical protein T20B5.2
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 16-Feb-2001
 C:Accession: T00746; C84834
 R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
 submitted to the EMBL Data Library, November 1997
 A:Description: Arabidopsis thaliana chromosome II BAC T20B5 genomic sequence.
 A:Reference number: Z14159
 A:Accession: T00746
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-945 <RND>
 A:Cross-References: EMBL:AC002409; NID:92623294; PID:92623296
 A:Experimental source: cultivar Columbia
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 Gues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083467
 A:Accession: C84834
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-945 <STO>
 A:Cross-References: GB:AE002093; NID:92623296; PIDN:AAB6442.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: Atg40820; T20B5.2
 A:Map position: 2
 A:introns: 111/3; 154/3; 185/3; 234/3; 262/3; 323/1; 362/1; 381/3; 425/3; 497/1; 670/1;

Query Match 83.3%; Score 35; DB 2; Length 945;

Best Local Similarity 85.7%; Pred. No. 6.1e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 MOPPLP 7
 |||||
 DB 806 MQPPPP 812

Search completed: August 8, 2002, 07:48:41
 Job time: 134 sec

KW Repeat; Cell wall; Glycoprotein; Signal; Structural protein;
 RM Hydroxylation. 1 20 POTENTIAL.
 FT SIGNAL 21 620 EXTERNSIN.
 FT CHAIN 70 73 H-A-P-P.
 FT REPEAT 148 151 H-A-P-P.
 FT DOMAIN 229 242 2 x 7 AA TANDEM REPEATS OF T-H-R-H-A-P-P.
 FT REPEAT 229 235 1.
 FT REPEAT 236 242 2.
 FT DOMAIN 205 620 CONSTRAINTS THE SER-PRO(4) REPEATS.
 FT DOMAIN 499 600 3 X APPROXIMATE TANDEM REPEATS.
 SQ SEQUENCE 620 AA; 65406 MW; 641DD2278AB28524 CRC64;

Query Match 88.1%; Score 37; DB 1; Length 620;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 DB 425 OPPPLP 430

RESULT 9

ID DREB_HUMAN STANDARD; PRT; 649 AA.

AC Q16643; 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Drebrin E.
 GN DBN1.
 OS Homo sapiens (Human).
 OC Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Osteoblast;
 RA Fisher L.W., McBride O.W., Filipula D., Ibaraki K., Young M.F.;
 RT "Human drebrin: cDNA sequence, mRNA tissue distribution and
 RT chromosomal localization.";
 RL Neurosci. Res. Commun. 14:35-42(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Fetal brain;
 RX MEDLINE-94030036; PubMed-8216329;
 RA Toda M., Shitao T., Minoshima S., Shimizu N., Toya S., Uyemura K.;
 RT "Molecular cloning of cDNA encoding human drebrin E and chromosomal
 RT mapping of its gene.";
 RL Biochem. Biophys. Res. Commun. 196:468-472(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Eye;
 RA Strusberg R.;
 RT Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: DREBRIN MIGHT PLAY SOME ROLE IN CELL MIGRATION.
 CC EXTENSION OF NEURONAL PROCESSES AND PLASTICITY OF DENDRITES,
 CC RESPECTIVELY. BINDS ACTIN.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: BRAIN NEURONS. ALSO FOUND IN THE HEART,
 CC PLACENTA, SKELETAL MUSCLE, KIDNEY AND PANCREAS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; D17530; BAA04480.1; -
 DR

DR EMBL; BC000283; AAH00283.1; -
 DR MIM; 126660; -
 DR InterPro; IPR002108; Cofilin_ADF.
 DR Pfam; PF00241; Cofilin_ADF; 1.
 DR SMART; SM00102; ADF; 1.
 KW Actin-binding; Brain; Neurone.
 SQ SEQUENCE 649 AA; 71425 MW; A1730E7C3FC32D88 CRC64;

Query Match 88.1%; Score 37; DB 1; Length 649;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 DB 363 OPPPLP 368

RESULT 10

ID AXN_DROME STANDARD; PRT; 745 AA.

AC Q9V407; 09YXC1; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Axin (axis inhibition protein) (daxin) (d-Axin).
 GN AXN OR CG7926.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 RN NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Embryo;
 RX MEDLINE-99174088; PubMed-10073940;
 RA Hamada F., Tomoyasu Y., Takatsu Y., Nakamura M., Nagai S.-I.,
 RA Suzuki A., Fujita F., Shibuya H., Toyoshima K., Ueno N., Akiyama T.;
 RT "Negative regulation of Wingless signaling by D-axin, a Drosophila
 RT homolog of axin.";
 RL Science 283:1739-1742(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ruel L., Anthopoulos N., Goncalves J., Manoukian A.S., Woodgett J.R.;
 RT "A Drosophila homolog of the axin gene is involved in the transduction
 RT of the wingless signal regulating the stability of the armadillo
 RT protein.";
 RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;

RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.G., Rogers J.H.C., Blazek R.G., Chame M., Pfeiffer B.D.,
 RA Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokov D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Chew S., Dahlke S., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
 RA Jalili M., Kalush F., Kapen G.H., Ke Z., Kienison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palczolo K., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relvert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Splier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskis R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT *The genome sequence of *Drosophila melanogaster*.
RT Science 287:2185-2195(2000).
RM (4)
RP FUNCTION.
RA MEDLINE-99387984; PubMed-10457025;
RA Willert K., Logan C.Y., Arcora A., Fish M., Nusse R.;
RT *A *Drosophila* Axin homolog, Daxin, inhibits Wnt signaling.*;
RT Development 126:4165-4173(1999).
CC -1- FUNCTION: INHIBITOR OF THE Wg SIGNALING PATHWAY. DOWN-REGULATES
CC BETA-CATENIN (ARMADILLO-ARM). PROBABLY FACILITATE THE
CC PHOSPHORYLATION OF BETA-CATENIN AND APC BY GSK-3B (ZESTE-WHITE
CC 3-ZW3).
CC -1- SUBUNIT: INTERACTS WITH ZW3 AND ARM. THE INTERACTION BETWEEN AXN
CC AND ARM OCCURS VIA THE ARMADILLO REPEATS CONTAINED IN ARM.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- DEVELOPMENTAL STAGE: UBIQUITOUSLY EXPRESSED THROUGHOUT THE
CC DEVELOPMENT.
CC -1- SIMILARITY: CONTAINS 1 RGS DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 DIX DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AF086811; AAD2486.1; -
CC EMBL; AF091813; AAF21293.1; -
CC EMBL; AE003772; AAF56993.1; -
CC FlyBase; FBgn0026597; Axn.
CC InterPro; IPR001158; DIX.
CC InterPro; IPR000342; RGS.
CC InterPro; IPR001849; PH.
CC Pfam; PF00778; DIX; 1.
CC ProDom; PD001580; RGS; 1.
CC ProDom; PD003639; DIX; 1.
CC SMART; SM00021; DAX; 1.
CC SMART; SM00315; RGS; 1.
CC PROSITE; PS01312; RGS; 1.
CC PROSITE; PS01312; RGS; 1.
CC Developmental protein.
CC
CC FT DOMAIN 54 172 RGS.
CC FT DOMAIN 640 646 POLY-SER.
CC FT DOMAIN 663 745
CC FT CONFLICT 454 454 DIX.
CC FT CONFLICT 644 645 R -> Q (IN REF. 1).
CC FT CONFLICT 745 745 MISSING (IN REF. 1).
CC SEQUENCE 745 AA; 81718 MW; 31A502528CEB84BA CRC64;

Query Match 88.1%; Score 37; DB 1; Length 745;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
DB 655 QPPPLP 660

RESULT 11
ID DYNL CAEEL STANDARD; PRT; 830 AA.
AC DYNL CAEEL
DT P39055;
DT 01-FEB-1995 (Rel. 31, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Dyanamin (EC 3.6.1.50).
GN DYN-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Felodieridae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-97439883; PubMed-9294229;
RA Clark S.G., Shurland D.L., Meyerowitz E.M., Bargmann C.I.,
RT van der Bliek A.M.;
RT *A dyanamin GTPase mutation causes a rapid and reversible temperature-
RT inducible locomotion defect in *C. elegans*.
RT Proc. Natl. Acad. Sci. U.S.A. 94:10438-10443(1997).
RN [2]
RP REVISIONS TO C-TERMINUS.
RC STRAIN-BRISTOL N2;
RA van der Bliek A.M.;
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RT -1- FUNCTION: MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN INVOLVED
RT IN PRODUCING MICROTUBULE BUNDLES AND ABLE TO BIND AND HYDROLYZE
RT GTP. MOST PROBABLY INVOLVED IN VESICULAR TRAFFICKING PROCESSES, IN
RT PARTICULAR ENDOCYTOSIS.
CC -1- CATALYTIC ACTIVITY: GTP + H(2)O = GDP + phosphate.
CC -1- SUBCELLULAR LOCATION: MICROTUBULE-ASSOCIATED.
CC -1- SIMILARITY: BELONGS TO THE DYANMIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; L29031; AAB72228.2; -
CC HSSP; 005193; DYN.
CC InterPro; IPR001401; Dyanamin.
CC InterPro; IPR000375; Dyanamin_central.
CC InterPro; IPR001310; GED.
CC InterPro; IPR001849; PH.
CC Pfam; PF00350; dyanamin_1.
CC Pfam; PF01031; dyanamin_2; 1.
CC Pfam; PF02212; GED; 1.
CC Pfam; PF00169; PH; 1.
CC PRINTS; PR00195; DYANMIN.
CC SMART; SM00053; DYN; 1.
CC SMART; SM00302; GED; 1.
CC SMART; SM00233; PH; 1.
CC PROSITE; PS00410; DYANMIN; 1.
CC PROSITE; PS00003; PH_DOMAIN; 1.
CC Hydrolyase; Motor protein; GTP-binding; Microtubules; Multigene family;
CC Endocytosis.
CC
CC FT NP_BIND 40 47 GTP (BY SIMILARITY).
CC FT NP_BIND 138 142 GTP (BY SIMILARITY).
CC FT NP_BIND 207 210 GTP (BY SIMILARITY).
CC FT DOMAIN 519 624 PH.
CC SEQUENCE 830 AA; 93348 MW; FC3D7106D079EDC5 CRC64;

Query Match 88.1%; Score 37; DB 1; Length 830;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
DB 747 OPPPLP 752

RESULT 12

524D_HUMAN STANDARD; PRT; 1032 AA.

AC 094855; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein transport protein Sec24D (SEC24-related protein D).
GN SEC24D OR KIA0755.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
[1]

SEQUENCE FROM N.A.

RA MEDLINE-99087487; PubMed-9872452;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XI.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL Dna Res. 5:277-286(1998).
[2]

SEQUENCE FROM N.A.

RA TISSUE-Pancreas;
RC MEDLINE-99262157; PubMed-10329445;
RA Tang B.L., Kausalya J., Low D.Y.H., Lock M.L., Hong N.;
RT "A family of mammalian proteins homologous to yeast Sec24p.";
RT Biochem. Biophys. Res. Commun. 258:679-684(1999).
CC -1- FUNCTION: COMPONENT OF THE COP1 COAT, THAT COVERS ER-DERIVED
CC VESICLES INVOLVED IN TRANSPORT FROM THE ENDOPLASMIC RETICULUM TO
CC THE GOLGI APPARATUS. COP1 ACTS IN THE CYTOSOL TO PROMOTE THE
CC TRANSPORT OF SECRETORY, PLASMA MEMBRANE, AND VACUOLAR PROTEINS
CC FROM THE ENDOPLASMIC RETICULUM TO THE GOLGI COMPLEX.
CC -1- SUBUNIT: COP1 IS COMPOSED OF AT LEAST FIVE PROTEINS: THE SEC23/24
CC COMPLEX, THE SEC13/31 COMPLEX AND SARL.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND PERINUCLEAR (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED, WITH HIGHER AMOUNTS IN
CC PLACENTA, PANCREAS, HEART AND LIVER.
CC -1- SIMILARITY: BELONGS TO THE SEC23/SEC24 FAMILY. SEC24 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AB018298; BAA34475.1; -;
DR EMBL: A010464; AAD28756.2; -;
KW Transport; Protein transport; Golgi stack; Endoplasmic reticulum;
KW Multigene family.
FT DOMAIN 363 388 ZINC FINGER-LIKE.
FT DOMAIN 10 293 PRO-RICH.
SQ SEQUENCE 1032 AA; 112999 MW; 1EAA80215EB979A7 CRC64;

Query Match 88.18; Score 37; DB 1; Length 1032;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
DB 207 OPPPLP 212

RESULT 13

TCF8_HUMAN STANDARD; PRT; 1124 AA.

AC P37275; Q13800; Q12924;
DT 01-OCT-1994 (Rel. 30, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor 8 (NFI-2-A zinc finger protein) (Negative
DE regulator of IL2).
GN TCF8 OR AREB6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
[1]

RA MEDLINE-94186507; PubMed-8138542;
RA Watanabe Y., Kawakami K., Hirayama Y., Nagano K.;
RT "Transcription factors positively and negatively regulating the Na,K-
RT ATPase alpha 1 subunit gene.";
RL J. Biochem. 114:849-855(1993).
[2]

SEQUENCE FROM N.A.

RA Bachman N.J., Scarpulla R.C.;
RT "A human zinc finger homeodomain protein homologous to the chicken
RT delta-crystallin enhancer binding protein, delta EPI.";
RL Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.
[3]

RA MEDLINE-92108424; PubMed-1640704;
RA Williams T.M., Woolsten D., Burtin J., Romano J., Bhargava R.,
RA Goddard A., Mellon M., Rauscher F.J. III, Kant J.A.;
RT "Identification of a zinc finger protein that inhibits IL-2 gene
RT expression.";
RL Science 254:1791-1794(1991).
CC -1- FUNCTION: INHIBITS INTERLEUKIN-2 (IL-2) GENE EXPRESSION. MAY BE
CC RESPONSIBLE FOR TRANSCRIPTIONAL REPRESSION OF THE IL-2 GENE.
CC ENHANCES OR REPASSES THE PROMOTER ACTIVITY OF THE AP1A1 GENE
CC DEPENDING ON THE QUANTITY OF CDNA AND ON THE CELL TYPE.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART AND SKELETAL MUSCLE, BUT
CC NOT IN LIVER, SPLEEN, OR PANCREAS.
CC -1- SIMILARITY: BELONGS TO DELTA-EPI/2FH-1 FAMILY OF TWO-HANDED ZINC
CC FINGER/HOMEOBOX PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D15050; BAA03646.1; -;
DR EMBL: U121170; AAA20602.1; -;
DR EMBL: M81699; -; NOT_ANNOTATED_CDS.
DR HSP: P08047; ISP2.
DR MIM: 189909; -;
DR InterPro: IPR001356; Homeobox.
DR InterPro: IPR000823; Znf-C2H2.
DR Pfam: PF00096; Zf-C2H2; 7.
DR PRINTS: PR00046; ZINCfinger.
DR SMART: SM00389; HOX; 1.
DR SMART: SM00355; Znf-C2H2; 7.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 5.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 6.
KW Transcription regulation; DNA-binding; Nuclear protein; Zinc-finger;
KW Homeobox; Repressor; Activator; Metal-binding; Repeat.
FT ZN_FING 170 193 C2H2-TYPE.
FT ZN_FING 200 222 C2H2-TYPE.
FT ZN_FING 240 262 C2H2-TYPE.
FT ZN_FING 268 292 C2H2-TYPE (ATYPICAL).
FT DNA_BIND 581 640 HOMEBOX-LIKE.

FT 2N_PING 904 926 C2H2-TYPE.
 FT 2N_PING 932 934 C2H2-TYPE.
 FT 2N_PING 960 981 C2H2-TYPE (ATYPICAL).
 FT DOMAIN 989 1124 GLU-RICH (ACIDIC).
 FT CONFLICT 420 420 V -> I (IN REF. 2).
 FT CONFLICT 609 609 E -> Q (IN REF. 3).
 FT CONFLICT 654 654 I -> T (IN REF. 2).
 FT CONFLICT 672 672 D -> H (IN REF. 3).
 FT CONFLICT 681 681 L -> S (IN REF. 3).
 SQ SEQUENCE 1124 AA; 124073 MW; 0A271ACC37C848D1 CRC64;

Query Match 88.1%; Score 37; DB 1; Length 1124;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OPPPLP 7
 DB 557 OPPPLP 562

RESULT 14
 ID FMN2_MOUSE STANDARD; PRT: 1567 AA.
 AC 09J104; 2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Formin 2.
 GN FMN2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE-Brain;
 RX MEDLINE=20243524; PubMed=10781961;
 RA Leader B., Leder P.;
 RT Formin-2, a novel formin homology protein of the cappuccino
 RT subfamily, is highly expressed in the developing and adult central
 RT nervous system.*
 RL Mech. Dev. 93:221-231(2000).

CC -1- TISSUE SPECIFICITY: EXPRESSED ALMOST EXCLUSIVELY IN THE DEVELOPING
 CC AND MATURE CENTRAL NERVOUS SYSTEM.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION BEGINS AT EMBRYONIC DAY 9. 5 IN
 CC THE DEVELOPING SPINAL CORD AND BRAIN STRUCTURES AND CONTINUES IN
 CC NEONATAL AND ADULT BRAIN STRUCTURES INCLUDING THE OLFACTORY BULB,
 CC CORTICAL THALAMUS, HIPPOCAMPUS AND CEREBELLUM.
 CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOLOG 1 (FH1) DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOLOG 2 (FH2) DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE FORMIN HOMOLOG FAMILY. CAPPUCCINO
 CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC -----
 CC EMBL: AF218940; AAF72883.1; -
 CC MGD: MG1:1859252; Fmn2.
 CC InterPro: IPR003104; FH2.
 CC InterPro: IPR001265; Formin.
 CC Pfam: PF02181; FH2.1.
 CC PRINTS: PR00828; FORMIN.
 CC SMART: SM00498; FH2.1.
 DR Developmental protein; Repeat; Coiled coil.
 KW DOMAIN 643 683 COILED COIL (POTENTIAL).
 FT DOMAIN 735 1113 FH1 (PRO-RICH).
 FT DOMAIN 919 1039 11 X 11 AA TANDEM REPEATS OF [MV]-G-I-P-

FT REPEAT 919 929 P-P-P-P-L-P-G.
 FT REPEAT 930 940 1.
 FT REPEAT 941 951 2.
 FT REPEAT 952 962 3.
 FT REPEAT 963 973 4.
 FT REPEAT 974 984 5.
 FT REPEAT 985 995 6.
 FT REPEAT 996 1006 7.
 FT REPEAT 1007 1017 8.
 FT REPEAT 1018 1028 9.
 FT REPEAT 1029 1039 10.
 FT REPEAT 1128 1132 11.
 FT DOMAIN 1408 1444 FH2.
 FT DOMAIN 48 55 COILED COIL (POTENTIAL).
 FT DOMAIN 202 207 POLY-GLN.
 FT DOMAIN 797 801 POLY-GLN.
 FT DOMAIN 861 864 POLY-PRO.
 FT DOMAIN 908 917 POLY-PRO.
 FT DOMAIN 922 928 POLY-PRO.
 FT DOMAIN 933 939 POLY-PRO.
 FT DOMAIN 944 950 POLY-PRO.
 FT DOMAIN 955 961 POLY-PRO.
 FT DOMAIN 966 972 POLY-PRO.
 FT DOMAIN 977 983 POLY-PRO.
 FT DOMAIN 988 994 POLY-PRO.
 FT DOMAIN 999 1005 POLY-PRO.
 FT DOMAIN 1010 1016 POLY-PRO.
 FT DOMAIN 1021 1027 POLY-PRO.
 FT DOMAIN 1032 1038 POLY-PRO.
 FT DOMAIN 1043 1046 POLY-PRO.
 FT DOMAIN 1054 1060 POLY-PRO.
 FT DOMAIN 1065 1072 POLY-PRO.
 FT DOMAIN 1077 1080 POLY-PRO.
 SQ SEQUENCE 1567 AA; 166268 MW; 8F273B1C88505944 CRC64;

Query Match 88.1%; Score 37; DB 1; Length 1567;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OPPPLP 7
 DB 898 OPPPLP 903

RESULT 15
 ID YAD4_YEAST STANDARD; PRT: 463 AA.
 AC P28003;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Hypothetical 53.1 kDa protein in PYK1-SN1 intergenic region.
 GN YAL034C OR FUM19.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_Taxid=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92260538; PubMed=1583694;
 RA Harris S.D., Cheng J., Pugh T.A., Pringle J.R.;
 RT Molecular analysis of Saccharomyces cerevisiae chromosome I. On the
 RT number of genes and the identification of essential genes using
 RT temperature-sensitive-lethal mutations.*
 RL J. Mol. Biol. 225:53-65(1992).

CC [2]
 CC SEQUENCE FROM N.A.
 CC STRAIN=S288C / AB972;
 CC MEDLINE=95249563; PubMed=7731988;
 CC Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortlin N.,
 CC Hall J., Ouellette B.F.F., Keng T., Barton A.B., Su Y., Davies C.K.,

RA Storms R.K.:
 RT "The nucleotide sequence of chromosome I from *Saccharomyces*
 RT *cerevisiae*.";
 RT Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813(1995).
 RN [3]
 RN RP REVISIONS.
 RA Vo D.T.:
 RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- MISCELLANEOUS: FUN19 IS A NON-ESSENTIAL GENE.
 CC -1- SIMILARITY: TO S.POMBE SPAC14C4.12C.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X62577; CAA44456.1; -
 CC EMBL: U12980; AAC04998.1; -
 CC DR PIR: S23410; S23410.
 CC DR SGD: S0002134; FUN19.
 CC KW Hypothetical protein.
 CC FT CONFLICT 305 305 T -> S (IN REF. 1)
 CC FT CONFLICT 416 463 VYRLKGLPFRRTDQKACRIDVKNKSLFQAPKYGWL
 CC FT ODSNFTKYL -> FIDQRCGRLEGPTPKRPGVSTLKH
 CC FT ODPYSLRLAGYRIRILRLSTYNTGCMHFKHGLLSNF
 CC FT LAFYLIHS (IN REF. 1).
 CC FT SEQUENCE 463 AA; 53107 MW; 420204B5562DBRDD CRC64;
 SO

Query Match 85.7%; Score 36; DB 1; Length 463;
 Best Local Similarity 85.7%; Pred. No. 95;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 Db 243 MTPPLP 249

RESULT 16
 WASL_BOVIN STANDARD; PRT; 505 AA.
 AC Q95107;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Neural Wiskott-Aldrich syndrome protein (N-WASP).
 GN WASL.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97050838; PubMed=8895577;
 RA Miki H., Miura K., Takenawa T.;
 RT "N-WASP, a novel actin-depolymerizing protein, regulates the cortical
 RT cytoskeletal rearrangement in a p12-dependent manner downstream of
 RT tyrosine kinases.";
 RL EMBO J. 15:5326-5335(1996).
 CC -1- FUNCTION: REGULATES ACTIN POLYMERIZATION BY STIMULATING THE ACTIN-
 CC NUCLEATING ACTIVITY OF THE ACTIN-RELATED PROTEIN 2/3 (ARP2/3)
 CC COMPLEX.
 CC -1- SUBUNIT: BINDS ACTIN AND ARP2/3 COMPLEX; INTERACTS WITH CDC42
 CC BINDS TO SH3 DOMAINS OF ASH/GRB2.
 CC -1- SIMILARITY: CONTAINS 1 GBD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 WH1 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 WH2 DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: D67066; BAA11082.1; -
 CC DR InterPro: IPR000095; PAK_box_P21_Rho_binding;
 CC DR InterPro: IPR000697; RanBP1_WASP.
 CC DR InterPro: IPR001960; WH1.
 CC DR InterPro: IPR003124; WH2.
 CC DR Pfam: PF00786; PBD; 1.
 CC DR Pfam: PF00568; WH1; 1.
 CC DR Pfam: PF02205; WH2; 2.
 CC DR SMART: SM00285; PBD; 1.
 CC DR SMART: SM00461; WH1; 1.
 CC DR SMART: SM00246; WH2; 2.
 CC DR PROSITE: PS50108; GBD; 1.
 CC KW Actin-binding; Repeat.
 CC FT DOMAIN 34 138 WH1.
 CC FT DOMAIN 203 222 GBD.
 CC FT DOMAIN 277 392 PRO-RICH.
 CC FT DOMAIN 405 422 WH2 1.
 CC FT DOMAIN 433 450 WH2 2.
 CC FT DOMAIN 486 505 ASP-RICH.
 CC FT SEQUENCE 505 AA; 54671 MW; 54B83B48F1CDB3B8 CRC64;
 SO

Query Match 85.7%; Score 36; DB 1; Length 505;
 Best Local Similarity 85.7%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 Db 342 MTPPLP 348

RESULT 17
 DVL2_XENLA STANDARD; PRT; 736 AA.
 AC P51142;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Segment polarity protein dishevelled homolog DVL-2 (Dishevelled-2)
 DE (DSH homolog 2) (Xdsr).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OC NCBI_Taxid=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=95324391; PubMed=7600981;
 RA Sokol S.Y., Klingensmith J., Perrimon N., Itoh K.;
 RT "Dorsalizing and neuralizing properties of Xdsr, a maternally
 RT expressed Xenopus homolog of dishevelled.";
 RL Development 121:1637-1647(1995).
 RN [2]
 RP ERRATUM.
 RA MEDLINE=96017659; PubMed=7588081;
 RX Sokol S.Y., Klingensmith J., Perrimon N., Itoh K.;
 RL Development 121:3487-3487(1995).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE SIGNAL TRANSDUCTION PATHWAY
 CC MEDIATED BY MULTIPLE WNT GENES. PLAYS A ROLE IN DORSAL AXIS
 CC FORMATION AND IN NEURAL INDUCTION.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS PROTEIN FOUND EQUALLY DISTRIBUTED
 CC IN BOTH ANIMAL-VEGETAL AND DORSAL-VENTRAL DIRECTIONS.
 CC -1- DEVELOPMENTAL STAGE: MATERNAL GENE DETECTED IN DIFFERENT

DEVELOPMENTAL STAGES BEING MOST ABUNDANT IN EGGS.

CC -1- SIMILARITY: BELONGS TO THE DSH FAMILY.

CC -1- SIMILARITY: CONTAINS 1 PDZ DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U31552; AAB00688.1; -

DR HSSP: 012923; 3PDZ.

DR InterPro: IPR000591; DEP.

DR InterPro: IPR001158; DIX.

DR InterPro: IPR003351; Dishevelled.

DR InterPro: IPR001478; PDZ.

DR Pfam: PF00610; DEP. 1.

DR Pfam: PF02377; Dishevelled; 1.

DR Pfam: PF00778; DIX; 1.

DR Pfam: PF00595; PDZ; 1.

DR ProDom: PD003639; DIX; 1.

DR SMART: SM00021; DAX; 1.

DR SMART: SM00049; DEP; 1.

DR SMART: SM00228; PDZ; 1.

DR PROSITE: PS50186; DEP; 1.

DR PROSITE: PS50106; PDZ; 1.

DR Developmental protein.

KW DOMAIN 99 113 POLY-PRO.

FT DOMAIN 222 227 POLY-ARG.

FT DOMAIN 254 326 PDZ.

FT DOMAIN 428 502 DEP.

FT DOMAIN 680 687 POLY-PRO.

FT SEQUENCE 736 AA; 79787 MW; AF6C9A1662DD7CEB CRC64;

Query Match Best Local Similarity 85.7%; Score 36; DB 1; Length 736;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOPPLP 7
1 |11111
Db 679 MPPPLP 685

RESULT 18

NC1_NEUCR STANDARD; PRT; 823 AA.

AC P20824; 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Phosphorus acquisition controlling protein.

GN NUC-1.

OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Sordariales; Sordariaceae; Neurospora.

OX NCBI_Taxid=5141;

RP SEQUENCE FROM N.A.

RC STRAIN=40-21;

RX MEDLINE=91042513; PubMed=2146493;

RA Kang S., Metzberg R.L.; a gene controlling phosphorus acquisition in Neurospora crassa.

RT Molecular analysis of nuc-1+, a gene controlling phosphorus acquisition in Neurospora crassa.

RL Mol. Cell. Biol. 10:5839-5848(1990).

CC -1- FUNCTION: FACTOR THAT ACTIVATES THE TRANSCRIPTION OF STRUCTURAL GENES FOR PHOSPHORUS ACQUISITION.

CC -1- SUBUNIT: BINDS DNA AS A DIMER.

CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF TRANSCRIPTION FACTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: M37700; AAA33603.1; -

DR PIR: A36378; A36378.

DR HSSP: F07270; 1A0A.

DR TRANSFAC: T01642; -

DR InterPro: IPR003015; HLH_MYC.

DR InterPro: IPR001092; HLH_dlm.

DR Pfam: PF00010; HLH; 1.

DR SMART: SM00353; HLH; 1.

DR PROSITE: PS00038; HELIX_LOOP_HELIX; 1.

KW DNA-binding; Transcription regulation; Nuclear protein; Activator.

FT DOMAIN 22 51 ASP-RICH (ACIDIC).

FT DOMAIN 101 220 GLN-RICH (INVOLVED IN TRANSCRIPTIONAL ACTIVATION) (POTENTIAL).

FT DOMAIN 434 556 PRO-RICH.

FT DOMAIN 468 562 INTERACTION WITH NEGATIVE REGULATORY FACTOR (POTENTIAL).

FT DOMAIN HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).

FT SEQUENCE 718 758 MW; 5E513ED98966E2F CRC64;

Query Match Best Local Similarity 85.7%; Score 36; DB 1; Length 823;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOPPLP 7
1 |11111
Db 433 MPPPLP 439

RESULT 19

OE56_NPVHZ STANDARD; PRT; 175 AA.

AC 010620; 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Occlusion-derived virus envelope protein E56 (ODV-E56) (ODV-6E) (Fragment).

GN ODVPE.

OS Heliothis zea nuclear polyhedrosis virus (HzSNPV) (Helicoverpa zea single nucleocapsid nuclear polyhedrosis virus).

OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;

OC Nucleopolyhedrovirus.

OX NCBI_Taxid=28290;

RP SEQUENCE FROM N.A.

RC STRAIN=ELKAR;

RA Le T.H., Wu T., Robertson A., Bulach D.M., Cowan P.J., Goodge K., Tribe D.E.; Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: STRUCTURAL PROTEIN THAT IS SPECIFIC FOR OCCCLUSION-DERIVED VIRUS (ODV) ENVELOPES BUT NOT OF BUDDED VIRUS (BV) (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC LOCALIZED TO THE ENVELOPE REGION OF PREOCCLUDED BUNDLES OF VIRIONS.

CC -1- SIMILARITY: TO CORRESPONDING PROTEIN IN OTHER BACULOVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC or send an email to license@lsb-sib.ch.
 CC EMBL: U67264; AAB54101.1; -
 DR Late protein; Envelope protein; Transmembrane.
 FT NON_TER 1
 FP TRANSMEM 139 159 POTENTIAL.
 SO SEQUENCE 175 AA; 19159 MW; 8159FBD84F091F1B CRC64;

Query Match 83.3%; Score 35; DB 1; Length 175;
 Best Local Similarity 85.7%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 DB 164 MOPPLP 170

RESULT 20
 PRS6_CAEEL STANDARD; PRT; 414 AA.
 AC P46502;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Probable 26S protease regulatory subunit 6B.
 GN F23P12.6
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Du 2;
 RL Submitted (AUG-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: THE 26S PROTEASE IS INVOLVED IN THE ATP-DEPENDENT
 CC DEGRADATION OF UNIDENTIFIED PROTEINS. THE REGULATORY (OR ATPASE)
 CC COMPLEX CONFERS ATP DEPENDENCY AND SUBSTRATE SPECIFICITY TO THE
 CC 26S COMPLEX (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (Potential).
 CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: U12965; AAA20608.1;
 DR WormPep; F23F12.6; CE01253.
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR003960; AAA_sub.
 DR InterPro: IPR003959; AAA_subfam.
 DR Pfam: PF00004; AAA; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00674; AAA; 1.
 KW Hypothetical protein; Proteasome; ATP-binding; Nuclear protein.
 FT NP_BIND 202 209 ATP (POTENTIAL).
 SO SEQUENCE 414 AA; 46358 MW; 8630AFB8A2C7F32 CRC64;

Query Match 83.3%; Score 35; DB 1; Length 414;
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 DB 20 LRPPPLP 26

RESULT 21
 MOTL_THEAQ STANDARD; PRT; 533 AA.
 ID MOTL_THEAQ
 AC P6082;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA mismatch repair protein mutL.
 GN MOTL.
 OS Thermus aquaticus.
 OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
 OX NCBI_TaxID=271;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-YT1;
 RA Yamamoto A., Biswas I., Hsieh P.;
 RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES IN
 CC DNA. IT IS REQUIRED FOR DAM-DEPENDENT METHYL-DIRECTED DNA MISMATCH
 CC REPAIR. MAY ACT AS A "MOLECULAR MATCHMAKER", A PROTEIN THAT
 CC PROMOTES THE FORMATION OF A STABLE COMPLEX BETWEEN TWO OR MORE
 CC DNA-BINDING PROTEINS IN AN ATP-DEPENDENT MANNER WITHOUT ITSELF
 CC BEING PART OF A FINAL EFFECTOR COMPLEX (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MOTL/HEX FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: U50453; AAB40601.1; -
 DR HSRP; P23367; 1BRN.
 DR InterPro: IPR002099; DNA_mis_repair.
 DR InterPro: IPR004359; HATPase_C.
 DR InterPro: IPR004359; His_Kin_sig.
 DR Pfam: PF01119; DNA_mis_repair; 1.
 DR Pfam: PF02518; HATPase_C; 1.
 DR SMART: SM00387; HATPase_C; 1.
 DR PROSITE: PS00058; DNA_MISMATCH_REPAIR_1; FALSE_NEG.
 KW DNA repair.
 SO SEQUENCE 533 AA; 58536 MW; 4442F9BB7A6CB8A4 CRC64;

Query Match 83.3%; Score 35; DB 1; Length 533;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 DB 329 LRPPPLP 335

RESULT 22
 YID6_YEAST STANDARD; PRT; 587 AA.
 ID YID6_YEAST
 AC P40535;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Hypothetical 65.3 kDa protein in NOT3-CKA1 intergenic region.
 GN YII036W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA Bairdell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,

DR PROSITE: PS50234; VMFA: 1.
 RM Photosynthesis: Bacterioloctophyll1 biosynthesis.
 FT DOMAIN 475 661 VMFA.
 FT DOMAIN 336 349 POLY-PRO.
 SQ SEQUENCE 666 AA; 72578 MW; SCB072BD72072BBS CRC64;

Query Match 83.3%; Score 35; DB 1; Length 666;
 Best Local Similarity 85.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 |||||
 DB 334 MOPPPP 340

RESULT 25
 1066_NPVAC STANDARD; PRT: 808 AA.
 AC P41467;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 94.0 kDa protein in POL-1REP3 intergenic region.
 OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolydnavirus.
 OX NCBI_TaxID=46015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6.
 RX MEDLINE=94303173; PubMed=8030224;
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Farber M., Possee R.D.;
 RT "The complete DNA sequence of Autographa californica nuclear
 polyhedrosis virus.";
 RT Virology 202:586-605(1994).
 RL -1- SIMILARITY: TO CORRESPONDING ORF IN OPMNPV AND LDMNPV.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L22858; AAA6696.1;
 DR Hypothetical protein.
 KW Hypothetical protein.
 FT DOMAIN 106 117 POLY-PRO.
 SQ SEQUENCE 808 AA; 93973 MW; 76A871D2B6633F8A CRC64;

Query Match 83.3%; Score 35; DB 1; Length 808;
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 |||||
 DB 104 MOPPPP 110

RESULT 26
 ATRX_MOUSE STANDARD; PRT: 2476 AA.
 ID ATRX_MOUSE
 AC Q61687;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Transcriptional regulator ATRX (X-linked nuclear protein)
 DE (Heterochromatin protein 2) (HP1 alpha-interacting protein) (HP1-BP38
 DE protein).
 DE ATRX OR XNP OR HP1BP2.
 GN ATRX OR XNP OR HP1BP2.
 OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98213653; PubMed=9345503;
 RA Picketts D.J., Tassan A.O., Higgs D.R., Gibbons R.J.;
 RT "Comparison of the human and murine ATRX gene identifies highly
 RT conserved, functionally important domains.";
 RT Mamm. Genome 9:400-403(1998).
 RN [2]
 RP SEQUENCE OF 325-1176 FROM N.A.
 RX MEDLINE=97133299; PubMed=8978696;
 RA Le Douarin B., Nielsen A.L., Garner J.-M., Ichinose H.,
 RA Jeanmougin F., Losson R., Chandon P.;
 RT "A possible involvement of TfR1 alpha and TfR1 beta in the epigenetic
 RT control of transcription by nuclear receptors.";
 RT EMBO J. 15:6701-6715(1996).
 RN [3]
 RP SUBCELLULAR LOCATION, AND ASSOCIATION WITH PERICENTROMERIC
 RP HETEROCHROMATIN.
 RX MEDLINE=20040663; PubMed=10570185;
 RA McQuell T.L., Gibbons R.J., Sutherland H., O'Rourke D.M.,
 RA Bickmore W.A., Pombo A., Turley H., Gatter K., Picketts D.J.,
 RA Buckle V.J., Chapman L., Rhodes D., Higgs D.R.;
 RT "Localization of a putative transcriptional regulator (ATRX) at
 RT pericentromeric heterochromatin and the short arms of acrocentric
 RT chromosomes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 96:13983-13988(1999).
 CC -1- FUNCTION: COULD BE A GLOBAL TRANSCRIPTIONAL REGULATOR. MODIFIES
 CC GENE EXPRESSION BY AFFECTING CHROMATIN.
 CC -1- SUBUNIT: PROBABLY BINDS EZH2. BINDS ANNEKIN V IN A CALCIUM AND
 CC PHOSPHATIDYLCHOLINE/PHOSPHATIDYLSELINE-DEPENDENT MANNER (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH PERICENTROMERIC
 CC HETEROCHROMATIN DURING INTERPHASE AND MITOSIS, PROBABLY BY
 CC INTERACTING WITH HP1.
 CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 PHD-TYPE ZINC FINGER.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF026032; AAC08741.1;
 DR EMBL; X99643; CA67962.1;
 DR MGD; MGI:103067; Xnp.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR000330; SNF2_N.
 DR InterPro: IPR001841; znf.fing.
 DR Pfam; PF00271; helicase_C; 1.
 DR Pfam; PF00176; SNF2_N; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC_C; 1.
 DR SMART; SM00184; RING; 1.
 DR DNA repair; Nuclear protein; DNA-binding; Helicase; ATP-binding;
 DR Zinc-finger.
 RN [4]
 RP PHD-TYPE.
 RX NP_BIND 1579 1586
 FT SITE 1704 1707 DEGH BOX.
 FT DOMAIN 735 738 POLY-SER.
 FT DOMAIN 735 738 POLY-SER.
 FT DOMAIN 1001 1004 POLY-GLU.
 FT DOMAIN 1130 1135 POLY-SER.
 FT DOMAIN 1182 1185 POLY-SER.
 FT DOMAIN 1238 1245 POLY-ASP.
 FT DOMAIN 1484 1487 POLY-GLU.
 FT DOMAIN 1924 1931 POLY-SER.

FT DOMAIN 2205 2208 POLY-LYS.
 FT DOMAIN 2245 2248 POLY-GLU.
 FT DOMAIN 2403 2408 POLY-GLN.
 SO SEQUENCE 2476 AA; 278601 MW; 90A442B790FC4FFAC CRC64;

Query Match 83.3%; Score 35; DB 1; Length 2476;
 Best Local Similarity 100.0%; Pred. No. 7.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M0PPL 6
 DB 2450 M0PPL 2455.

RESULT 27
 ATRX_HUMAN STANDARD; PRT; 2492 AA.
 ID ATRX_HUMAN P51068; Q15886; Q9NTS3; Q9H0Z1;
 AC P46100; P51068; Q15886; Q9NTS3; Q9H0Z1;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Transcriptional regulator ATRX (X-linked helicase II) (X-linked
 nuclear protein) (XNP) (Znf-HX).
 GN ATRX OR RAD54L OR XH2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5), AND VARIANT S-1860,
 RP MEDLINE-97123494; PubMed-8968741;
 RA MEDLINE-97123494; PubMed-8968741;
 RA Picketts D.J., Higgs D.R., Bachoo S., Blake D.J., Quarrell O.W.J.,
 RA Gibbons R.J.,
 RT "ATRX encodes a novel member of the SNF2 family of proteins: mutations
 RT point to a common mechanism underlying the ATR-X syndrome.";
 RT Hum. Mol. Genet. 5:1899-1907(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).
 RP MEDLINE-97386582; PubMed-9244431;
 RA Villard L., Lossi A.-M., Cardoso C., Proud V., Chlaroni P.,
 RA Colleaux L., Schwartz C., Fontes M.;
 RT "Determination of the genomic structure of the XNP/ATRX gene encoding
 RT a potential zinc finger helicase.";
 RT Genomics 43:149-155(1997).
 RN [3]
 RP SEQUENCE OF 860-2492 FROM N.A.
 RP MEDLINE-95179111; PubMed-7874112;
 RA Stayton C.L., Dabovic B., Gulisano M., Gecz J., Broccoli V.,
 RA Giovannazzi S., Bossolasco M., Monaco L., Rastan S., Boncinelli E.,
 RA Bianchi M.E., Gonzalez G.G.;
 RT "Cloning and characterization of a new human Xq13 gene, encoding a
 RT putative helicase.";
 RT Hum. Mol. Genet. 3:1957-1964(1994).
 RN [4]
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RP MEDLINE-94214473; PubMed-8162050;
 RA Gecz J., Pollard H., Gonzalez G., Villard L., Stayton C.L.,
 RA Millaesau P., Khrestchatsky M., Fontes M.;
 RT "Cloning and expression of the murine homologue of a putative human
 RT X-linked nuclear protein gene closely linked to PK1 in Xq13.3.";
 RT Hum. Mol. Genet. 3:39-44(1994).
 RN [5]
 RP SEQUENCE OF 2401-2492 FROM N.A., AND VARIANTS ATR-X.
 RP MEDLINE-95211835; PubMed-7697714;
 RA Gibbons R.J., Picketts D.J., Villard L., Higgs D.R.;
 RT "Mutations in a putative global transcriptional regulator cause X-
 RT linked mental retardation with alpha-thalassemia (ATR-X syndrome).";
 RT Cell 80:837-845(1995).
 RN [6]
 RP SEQUENCE OF 1375-2492 FROM N.A.
 RP Pearce A., Chapman J.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP E2H2 BINDING.
 RX MEDLINE-98167853; PubMed-9499421;
 RA Cardoso C., Timsit S., Villard L., Khrestchatsky M., Fontes M.,
 RA Colleaux L.;
 RT "Specific interaction between the XNP/ATR-X gene product and the SET
 RT domain of the human E2H2 protein.";
 RT Hum. Mol. Genet. 7:679-684(1998).
 RN [8]
 RP SUBCELLULAR LOCATION, AND ASSOCIATION WITH PERICENTROMERIC
 RP HETEROCHROMATIN.
 RX MEDLINE-20040663; PubMed-10570185;
 RA McDowell T.L., Gibbons R.J., Sutherland H., O'Rourke D.M.,
 RA Bickmore W.A., Pombo A., Turley H., Gatter K., Picketts D.J.,
 RA Buckle V.J., Chapman L., Rhodes D., Higgs D.R.;
 RT "Localization of a putative transcriptional regulator (ATRX) at
 RT pericentromeric heterochromatin and the short arms of acrocentric
 RT chromosomes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 96:13983-13988(1999).
 RN [9]
 RP DISEASE.
 RX MEDLINE-20213147; PubMed-10751095;
 RA Villard L., Fontes M., Ades L.C., Gecz J.;
 RT "Identification of a mutation in the XNP/ATR-X gene in a family
 RT reported as Smith-Fineman-Myers syndrome.";
 RT Am. J. Med. Genet. 91:83-85(2000).
 RN [10]
 RP VARIANT ATR-X SFR-1713.
 RX MEDLINE-97196774; PubMed-9043863;
 RA Villard L., Lacombe D., Fontes M.;
 RT "A point mutation in the XNP gene, associated with an ATR-X phenotype
 RT without alpha-thalassemia.";
 RT Eur. J. Hum. Genet. 4:316-320(1996).
 RN [11]
 RP VARIANT JM GLN-2131.
 RX MEDLINE-96224392; PubMed-8630485;
 RA Villard L., Gecz J., Mattei J.-F., Fontes M., Saugier-Verber P.,
 RA Munnich A., Lyonnet S.;
 RT "XNP mutation in a large family with Juberger-Marsidi syndrome.";
 RT Nat. Genet. 12:359-360(1996).
 RN [12]
 RP VARIANTS ATR-X.
 RX MEDLINE-97467722; PubMed-9326931;
 RA Gibbons R.J., Bachoo S., Picketts D.J., Afflms S., Azenbauer B.,
 RA Bergoffen J., Berry S.A., Dahl N., Fryer A., Keppler K., Kurosawa K.,
 RA Levin M.L., Masuno M., Neri G., Pierpont M.E., Stanley S.F.,
 RA Higgs D.R.;
 RT "Mutations in transcriptional regulator ATRX establish the functional
 RT significance of a PHD-like domain.";
 RT Nat. Genet. 17:146-148(1997).
 RN [13]
 RP VARIANT ATR-X LEU-246.
 RX MEDLINE-20123062; PubMed-10660327;
 RA Ficheria M., Romano C., Castiglita L., Falla P., Ruberto C., Amata S.,
 RA Greco D., Cardoso C., Fontes M., Ragusa A.;
 RT "New mutations in XNP/ATR-X gene: a further contribution to
 RT genotype/phenotype relationship in ATR/X syndrome.";
 RT Hum. Mutat. 12:214-214(1998).
 RN [14]
 RP VARIANT SHS LYS-1742.
 RX MEDLINE-99347960; PubMed-10417298;
 RA Lossi A.-M., Millan J.M., Villard L., Orellana C., Cardoso C.,
 RA Prieto F., Fontes M., Martinez F.;
 RT "Mutation of the XNP/ATR-X gene in a family with severe mental
 RT retardation, spastic paraplegia and skewed pattern of X inactivation:
 RT demonstration that the mutation is involved in the inactivation
 RT bias.";
 RT Am. J. Hum. Genet. 65:558-562(1999).
 RN [15]
 RP VARIANT CMS THR-2050.
 RX MEDLINE-99326061; PubMed-10398237;
 RA Abidi F., Schwartz C.E., Carpenter N.J., Villard L., Fontes M.,

RA Curtis M.;
 RT "Carpenter-Haziri syndrome results from a mutation in XNP."
 RL Am. J. Med. Genet. 85:249-251(1999).
 RN [16]
 RP VARIANTS ATR-X E-175; 178-V-R-198 DEL; S-190; P-219; L-245 AND C-249.
 RX MEDLINE=99219535; PubMed=10204841;
 RA Villard L., Bonino M.-C., Abidi F., Ragusa A., Belouange J.,
 Ralossi A.-M., Seaver L., Bonnefont J.-P., Romano C., Fichera M.,
 RA Lacombe D., Hanauer A., Philip N., Schwartz C.E., Fontes M.;
 RT "Evaluation of a mutation screening strategy for sporadic cases of
 RT ATR-X syndrome."
 RL J. Med. Genet. 36:183-186(1999).
 RN [17]
 RP VARIANTS ATR-X S-179; L-190; I-194; C-246; F-1552; S-1645 AND C-1847.
 RX MEDLINE=20451413; PubMed=10995512;
 RA Wada T., Kubota T., Fukushima Y., Saitoh S.;
 RT "Molecular genetic study of Japanese patients with X-linked
 RT alpha-thalassemia/mental retardation syndrome (ATR-X)."
 RL Am. J. Med. Genet. 94:242-248(2000).
 CC -1 FUNCTION: COULD BE A GLOBAL TRANSCRIPTIONAL REGULATOR. MODIFIES
 CC GENE EXPRESSION BY AFFECTING CHROMATIN. MAY BE INVOLVED IN
 CC BRAIN DEVELOPMENT AND FACIAL MORPHOGENESIS.
 CC -1 SUBUNIT: PROBABLY BINDS EZH2. BINDS ANNEKIN V IN A CALCIUM AND
 CC PHOSPHATIDYLCHOLINE/PHOSPHATIDYLSELINE-DEPENDENT MANNER (BY
 CC similarity).
 CC -1 SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH PERICENTROMERIC
 CC HETEROCHROMATIN DURING INTERPHASE AND MITOSIS, PROBABLY BY
 CC INTERACTING WITH HPI.
 CC -1 ALTERNATIVE PRODUCTS: 5 ISOFORMS, 1, 2, 3, 4 (SHOWN HERE) AND 5;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1 TISSUE SPECIFICITY: UBIQUITOUS.
 CC -1 DISEASE: DEFECTS IN ATRX ARE THE CAUSE OF X-LINKED ALPHA-
 CC THALASSEMIA/MENTAL RETARDATION SYNDROME (ALSO KNOWN AS ATR-X
 CC SYNDROME). ATRX IS AN X-LINKED DISORDER COMPRISING SEVERE
 CC PSYCHOMOTOR RETARDATION, FACIAL DYSMORPHISM, URGENTAL,
 CC ABNORMALITIES, AND ALPHA-THALASSEMIA. AN ESSENTIAL PHENOTYPIC
 CC TRAIT ARE HEMOGLOBIN H ERYTHROCYTE INCLUSIONS.
 CC -1 DISEASE: DEFECTS IN ATRX ARE THE CAUSE OF CARPENTER-WIZIRI
 CC SYNDROME (CMS), AN X-LINKED RECESSIVE CONDITION CHARACTERIZED BY
 CC MODERATE MENTAL RETARDATION, SHORT STATURE, BRACHYDACTYL WITH
 CC EXCESSIVE SKIN CREASES, AND WIDENING OF THE KNUCKLES.
 CC -1 DISEASE: DEFECTS IN ATRX ARE THE CAUSE OF JUBERG-MARSDI SYNDROME
 CC (JM). JM IS A RARE X-LINKED RECESSIVE DISEASE CHARACTERIZED BY
 CC SEVERE MENTAL RETARDATION, GROWTH FAILURE, SENSORINEURAL DEAFNESS,
 CC MICROGENITALISM AND EARLY DEATH.
 CC -1 DISEASE: DEFECTS IN ATRX ARE A CAUSE OF SMITH-FINEMAN-MYERS
 CC RETARDATION, MICROCEPHALY, GROWTH FAILURE, FACIAL ANOMALIES AND
 CC BILATERAL CRYPTORCHIDISM. DUE TO THE CLINICAL OVERLAP WITH ATR-X
 CC SYNDROME, SOME PATIENTS ORIGINALLY DIAGNOSED AS HAVING SFM, MIGHT
 CC BE AFFECTED BY A VARIANT OF ATR-X SYNDROME WHICH LACK HEMOGLOBIN H
 CC INCLUSIONS.
 CC -1 DISEASE: DEFECTS IN ATRX ARE THE CAUSE OF SUTHERLAND-HAAN X-LINKED
 CC MENTAL RETARDATION SYNDROME (SHS). IT IS CHARACTERIZED BY SEVERE
 CC MENTAL RETARDATION WITH SPASTIC PARAPLEGIA, MICROCEPHALY, SHORT
 CC STATURE AND CRYPTORCHIDISM.
 CC -1 SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
 CC -1 SIMILARITY: CONTAINS 1 PHD-TYPE ZINC FINGER.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U72937; AAB49970.2; -
 CC EMBL: U72938; AAB49971.2; -
 CC EMBL: U72935; AAB40698.1; -
 CC EMBL: U72904; AAB40698.1; JOINED.
 DR
 Query Match 83.3%; Score 35; DB 1; Length 2492;

Best Local Similarity 100.0%; Pred. No. 7, 5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MOPPL 6
 Db 2467 MOPPL 2472
 ID y879 MYCTU STANDARD; PRT; 91 AA.
 AC 010541;
 DT 01-OCT-1996 (rel. 34, last sequence update)
 DT 01-OCT-1996 (rel. 34, last sequence update)
 DE 16-OCT-2001 (rel. 40, last annotation update)
 GN Hypothetical 9.5 kDa protein RV0879c.
 OS MYC0879C OR MF0902 OR MYCY31.07C.
 OC Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OC NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skellern S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CCC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwin M.L., Haft D., Salzman S.L.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishel W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: 273101; CA97386.1; -
 CC EMBL: AE006977; AAK45144.1; -
 CC DR TIGR: MT0902;
 CC TubercuList; RV0879c;
 DR Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 22 42
 FT TRANSMEM 53 73
 FT FT 71 82
 FT DOMAIN 71 82
 FT SEQUENCE 91 AA; 9512 MW; FDD6D8E9D162E033 CRC64;
 SO
 Query Match 81.0%; Score 34; DB 1; Length 91;
 Best Local Similarity 83.3%; Pred. No. 35;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 OPPPL 7
 :|||||

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2002, 07:47:27 ; Search time 10.47 Seconds

(Without alignments)
25.887 Million cell updates/sec

Title: US-09-641-801-1

Perfect score: 42

Sequence: 1 MOPPPP 7

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Listing first 150 summaries

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	92.9	222	1	COLI_THIOB
2	39	92.9	701	1	GGL_HUMAN
3	38	90.5	536	1	ARP_ARATH
4	37	88.1	167	1	SERO_GALME
5	37	88.1	211	1	RGM1_YEAST
6	37	88.1	278	1	FASL_RAT
7	37	88.1	483	1	BIAR_MELGA
8	37	88.1	620	1	EXTN_TOBAC
9	37	88.1	649	1	DREB_HUMAN
10	37	88.1	745	1	AXN_DROME
11	37	88.1	830	1	DYNI_CAEBL
12	37	88.1	1032	1	S24D_HUMAN
13	37	88.1	1124	1	TCF8_HUMAN
14	37	88.1	1567	1	FMN2_MOUSE
15	36	85.7	463	1	YAD4_YEAST
16	36	85.7	505	1	WASL_BOVIN
17	36	85.7	736	1	DVL2_XENLA
18	36	85.7	823	1	NTC1_NEOCR
19	35	83.3	175	1	OE56_NPVHZ
20	35	83.3	414	1	PRS6_CAEBL
21	35	83.3	533	1	MOTL_XENLA
22	35	83.3	587	1	YID6_YEAST
23	35	83.3	588	1	DAB1_MOUSE
24	35	83.3	666	1	BCHD_HELMO
25	35	83.3	808	1	Y066_NPVAC
26	35	83.3	2476	1	ATRX_MOUSE
27	35	83.3	2492	1	ATRX_HUMAN
28	34	81.0	91	1	Y879_MYCTU
29	34	81.0	193	1	PGC1_PIG
30	34	81.0	194	1	PGC1_HUMAN
31	34	81.0	194	1	PGC1_MOUSE
32	34	81.0	194	1	PGC1_RAT
33	34	81.0	202	1	SODF_NICPL

34	34	81.0	364	1	VT91_CAEBL	P48458 caenorhabd
35	34	81.0	410	1	AUP1_MOUSE	P70295 mus musculu
36	34	81.0	424	1	S3B4_HUMAN	Q15427 homo sapien
37	34	81.0	472	1	Y087_MYCTU	O53304 mycobacteri
38	34	81.0	509	1	SOX9_PIG	O18896 sus scrofa
39	34	81.0	555	1	DAB1_HUMAN	O75553 homo sapien
40	34	81.0	555	1	DAB1_MOUSE	O9B9X5 macaca fasc
41	34	81.0	652	1	FXO1_MOUSE	O9T1E0 mus musculu
42	34	81.0	718	1	AT12_HSV1	P10230 herpes slmp
43	34	81.0	992	1	EBN6_EBV	P03204 Epstein-Bar
44	34	81.0	1182	1	PC2_MOUSE	O35595 mus musculu
45	34	81.0	1255	1	PER2_HUMAN	O15055 homo sapien
46	34	81.0	1381	1	YBE7_YEAST	P34216 saccharomyc
47	34	81.0	1386	1	ZAP3_MOUSE	O9T017 mus musculu
48	34	81.0	1634	1	PK3B_HUMAN	O00750 homo sapien
49	33	78.6	282	1	ATF5_HUMAN	O9Y2D1 homo sapien
50	33	78.6	299	1	DHP1_HUMAN	O92874 homo sapien
51	33	78.6	322	1	UL15_HCVVA	P16844 human cytom
52	33	78.6	389	1	NDPP_MOUSE	O03173 mus musculu
53	33	78.6	447	1	AC11_CAEBL	P34429 caenorhabd
54	33	78.6	474	1	DNB2_ADEA1	P11807 human adeno
55	33	78.6	520	1	CET1_CANAL	O93803 candida alb
56	33	78.6	521	1	GAG_HV2CA	P24106 human Immun
57	33	78.6	526	1	SOA2_CERAE	O77159 cercopithec
58	33	78.6	679	1	PAN3_YEAST	P36102 saccharomyc
59	33	78.6	724	1	CCT1_MOUSE	O9QWV9 mus musculu
60	33	78.6	726	1	AD20_HUMAN	O43506 homo sapien
61	33	78.6	761	1	CCT1_HUMAN	O60563 homo sapien
62	33	78.6	748	1	CLAT_HUMAN	P28329 homo sapien
63	33	78.6	902	1	NFC4_HUMAN	O14934 homo sapien
64	33	78.6	923	1	SRR2_MOUSE	O9V712 mus musculu
65	33	78.6	1203	1	PRC2_HUMAN	O9T6C5 homo sapien
66	33	78.6	1387	1	RSCC_RAT	O08774 mus musculu
67	33	78.6	1807	1	TSO2_HUMAN	P49510 homo sapien
68	33	78.6	1822	1	ZAP3_HUMAN	O49750 mus musculu
69	33	78.6	2346	1	COAL_BOVIN	O9T553 bos taurus
70	33	78.6	2346	1	COAL_SHEEP	O28559 ovis aries
71	33	78.6	2715	1	TRX2_HUMAN	O9M066 homo sapien
72	32	76.2	74	1	UL11_HSVB	P28982 equine herp
73	32	76.2	122	1	SYN2_MOUSE	O91017 pseudomonas
74	32	76.2	126	1	NG3_DROME	O64332 mus musculu
75	32	76.2	165	1	NG3_DROME	P40140 drosophila
76	32	76.2	169	1	NAPB_ALCEU	P39186 alcaligenes
77	32	76.2	170	1	VEAR_HCVVA	P09694 human cytom
78	32	76.2	183	1	MGR1_MOUSE	P97772 mus musculu
79	32	76.2	198	1	LMBV_CHICK	O01636 gallus gall
80	32	76.2	211	1	YNO3_CAEBL	O03607 caenorhabd
81	32	76.2	226	1	GSPB_AERAY	P45755 aeromonas h
82	32	76.2	235	1	GAG_HVAB	P03333 abelson mur
83	32	76.2	255	1	AP1_ARATH	P35631 caenorhabd
84	32	76.2	256	1	HM34_CAEBL	O94165 caenorhabd
85	32	76.2	262	1	BCCP_SOYBN	O42783 glycine max
86	32	76.2	274	1	MEPA_ECOLI	P14007 escherichia
87	32	76.2	279	1	HRP5_HUMAN	O96K08 homo sapien
88	32	76.2	281	1	FASL_HUMAN	P48023 homo sapien
89	32	76.2	285	1	DXG2_CAEBL	O46552 caenorhabd
90	32	76.2	287	1	DLX3_HUMAN	O60479 homo sapien
91	32	76.2	287	1	DLX3_MOUSE	O64205 mus musculu
92	32	76.2	288	1	SMN_MOUSE	P97801 mus musculu
93	32	76.2	289	1	DHPS_SYNY3	P73248 synechocyst
94	32	76.2	290	1	TRX2_MOUSE	O08550 mus musculu
95	32	76.2	291	1	HX1M_HUMAN	O43711 homo sapien
96	32	76.2	291	1	HX1M_MOUSE	O55144 mus musculu
97	32	76.2	297	1	PUR7_MYCTU	O05966 mycobacteri
98	32	76.2	314	1	Y009_HUMAN	O15390 homo sapien
99	32	76.2	329	1	AAPC_PENCL	O40784 pennisetum
100	32	76.2	331	1	PEIB_COLGL	O59939 colletotric
101	32	76.2	331	1	PUR7_ARCFU	O28996 archoaeoglob
102	32	76.2	338	1	POSB_HUMAN	P53539 homo sapien
103	32	76.2	338	1	POSB_MOUSE	P13346 mus musculu
104	32	76.2	340	1	GXX2_CHICK	O42230 gallus gall
105	32	76.2	340	1	UL20_HCVVA	P16758 human cytom
106	32	76.2	356	1	HXB2_HUMAN	P16552 homo sapien

107	32	76.2	361	1	KITH_HSVSM	P04408 herpesvirus
108	32	76.2	363	1	PACI_HSHPO	P22192 schizosacch
109	32	76.2	363	1	TOBI_MOUSE	O61471 mus musculu
110	32	76.2	367	1	ND2_CHICK	P10360 gallus galli
111	32	76.2	382	1	ND2_RAT	O63689 rattus norv
112	32	76.2	382	1	PRLP_HUMAN	P51888 homo sapien
113	32	76.2	387	1	DADR_MOUSE	P51436 mus musculu
114	32	76.2	399	1	ALX4_MOUSE	O31137 mus musculu
115	32	76.2	408	1	G3PT_HUMAN	O41556 homo sapien
116	32	76.2	408	1	HBB_BOVIN	P42917 bos taurus
117	32	76.2	411	1	PAL1_HUMAN	O2um63 homo sapien
118	32	76.2	412	1	ALF_PETRY	O22621 petunia hyb
119	32	76.2	413	1	E2F4_HUMAN	O16254 homo sapien
120	32	76.2	413	1	HXD3_CHICK	O93353 gallus galli
121	32	76.2	415	1	KCCS_MALDO	O07250 malus domes
122	32	76.2	416	1	KCRU_BOVIN	O91848 bos taurus
123	32	76.2	419	1	KCRS_CHICK	P11009 gallus galli
124	32	76.2	419	1	KCRS_HUMAN	P17540 homo sapien
125	32	76.2	419	1	KCRS_RABIT	O77814 oryctolagus
126	32	76.2	419	1	KCRS_RAT	P09605 rattus norv
127	32	76.2	428	1	GRAP_BOVIN	O28115 bos taurus
128	32	76.2	431	1	PKNA_MYCTU	P71585 mycobacteri
129	32	76.2	432	1	GEAP_HUMAN	P14136 homo sapien
130	32	76.2	433	1	TEF5_CHICK	O90701 gallus galli
131	32	76.2	434	1	CXAX_HUMAN	O9y6h8 homo sapien
132	32	76.2	434	1	HPED_HOYU	O48604 hordeum vul
133	32	76.2	435	1	KICH_MOUSE	O54804 mus musculu
134	32	76.2	437	1	PKNA_MYCTU	P54743 mycobacteri
135	32	76.2	453	1	GAAG_HUMAN	O16445 homo sapien
136	32	76.2	453	1	KICH_RAT	O01134 rattus norv
137	32	76.2	456	1	KICH_HUMAN	P035790 homo sapien
138	32	76.2	458	1	RRA_NOTVI	P18514 notophthalm
139	32	76.2	458	1	RRA_XENLA	P18126 xenopus lae
140	32	76.2	460	1	RRA_CHICK	O90966 gallus galli
141	32	76.2	462	1	RRA_HUMAN	P10276 homo sapien
142	32	76.2	462	1	RRA_MOUSE	P11416 mus musculu
143	32	76.2	471	1	RB97_DROME	O02123 drosophila
144	32	76.2	477	1	CAP2_HUMAN	P40426 homo sapien
145	32	76.2	480	1	FXGB_RAT	O00939 rattus norv
146	32	76.2	481	1	FXGB_MOUSE	O60967 mus musculu
147	32	76.2	487	1	EBN2_EBV	P12378 Epstein-Bar
148	32	76.2	488	1	MM11_HUMAN	P24347 homo sapien
149	32	76.2	490	1	TA53_TREDE	P18164 treponema d
150	32	76.2	491	1	CPBS_RABIT	P12789 oryctolagus

ALIGNMENTS

RESULT 1
COLI_THUOB STANDARD: PRT: 222 AA.
AC 09YGR2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Corticotropin-lipotropin precursor (Pro-opiomelanocortin) (POMC)
DE [Contains: Corticotropin (Adrenocorticotrophic hormone) (ACTH);
DE Melanotropin alpha (Alpha-MSH); Corticotropin-like intermediate
DE peptide (CLIP); Lipotropin beta (Beta-LPH); Lipotropin gamma (Gamma-
DE LPH); Melanotropin beta (Beta-MSH); Beta-endorphin; Met-enkephalin].
OS Thunus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Plutary;
RA Amemiya Y., Takahashi A., Kawachi H.;
RT "Yuna piciopiomelanocortin cDNA."
RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

CC -1- PWM: SPECIFIC ENZYMAIC CLEAVAGES AT PAIRED BASIC RESIDUES YIELD
CC THE DIFFERENT ACTIVE PEPTIDES.
CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB020971; BAA35125.1; -
CC InterPro: IPR001941; Melanocortin_ACTH.
CC Pfam: PF00976; ACTH_domain.1.
CC PRINTS: PR00383; MELANOCORTIN.
CC PRODOM: PD003250; Melanocortin_ACTH; 1.
CC Endorphin; Hormone; Cleavage on pair of basic residues; Signal.
CC SIGNAL 1
CC PEPTIDE 93 132
CC PEPTIDE 93 107
CC PEPTIDE 111 132
CC PEPTIDE 136 222
CC PEPTIDE 136 188
CC PEPTIDE 172 188
CC PEPTIDE 191 222
CC PEPTIDE 191 195
CC PEPTIDE 71 75
CC DOMAIN 80 87
CC POLY-SER.
CC POLY-SER.
CC SEQUENCE 222 AA; 24970 MM; 7868C71306360B70 CRC64;

Query Match 92.9%; Score 39; DB 1; Length 222;
Best local similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDPPLP 7
:|||||
Db 62 IQPPLP 68
RESULT 2
CGI_HUMAN STANDARD: PRT: 701 AA.
ID CGI_HUMAN
AC 013495;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE CGI protein (F18).
DE CXORF6 OR CGI.
GN CXORF6 OR CGI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96225444; PubMed=8640223;
RA Laporte J., Hu L.-J., Kretz C., Mandel J.-L., Kioschis P., Coy J.,
RA Klauk S.M., Poutiska A., Dahl N.;
RT "A gene mutated in X-linked myotubular myopathy defines a new
RT putative tyrosine phosphatase family conserved in yeast."
RT Nat. Genet. 13:175-182(1996).
CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN SKELETAL MUSCLE.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U46023; AAC50551.1; -

```

DR  MIN: 300120;
FT  DOMAIN 275 286 POLY-PRO.
FT  DOMAIN 419 429 POLY-GLN.
FT  DOMAIN 522 533 POLY-GLN.
SQ  SEQUENCE 701 AA; 74475 MW; ADIC18EAF6D34E CRC64;

Query Match      92.9%; Score 39; DB 1; Length 701;
Best Local Similarity 85.7%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  1 MOPPLP 7
DB  273 LQPPPLP 279

RESULT 3
ARP_ARATH STANDARD; PRT; 536 AA.
AC  P45951;
DC  01-NOV-1995 (Rel. 32, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DE  01-MAR-2002 (Rel. 41, Last annotation update)
DE  Apurinic endonuclease-redox protein (DNA-apurinic or apyrimidinic
site) lyase (EC 4.2.99.18).
GN  ARP OR REP OR AT2G41460 OR T26U13.5.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eustosids II; Brassicales; Brassicaceae; Arabidopsids.
OX  NCBI_TaxID=3702;
[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=CV. COLUMBIA;
RC  MEDLINE=20083487; PubMed=10617197;
RA  Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA  Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldlyum T.V.,
RA  Buell C.R., Ketchum K.A., Lee J.C., Renning C.M., Koo H.L.,
RA  Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Umayam L.,
RA  Taitell L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA  Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA  Niemman M.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA  Venter J.C.;
RT  *Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT  thaliana.
RL  Nature 402:761-768(1999).
[2]
RN  SEQUENCE OF 10-536 FROM N.A.
RP  STRAIN=CV. COLUMBIA; TISSUE=Callus;
RC  MEDLINE=94211851; PubMed=7512729;
RA  Babychuk E., Kuschnit S., van Montagu M., Inze D.;
RA  "The Arabidopsis thaliana apurinic endonuclease Atp reduces human
RA  transcription factors fos and Jun.";
RT  Proc. Natl. Acad. Sci. U.S.A. 91:3299-3303(1994).
CC  -1- FUNCTION: REPAIRS OXIDATIVE DNA DAMAGES, SEEMS ALSO TO ACT AS A
CC  REOX FACTOR. IS MULTIFUNCTIONAL AND MAY BE INVOLVED BOTH IN DNA
CC  REPAIR AND IN THE RESOLUTION OF TRANSCRIPTION.
CC  -1- CATALYTIC ACTIVITY: The C-O-P bond 3' to the apurinic or
CC  apyrimidinic site in DNA is broken by a beta-elimination reaction,
CC  leaving a 3'-terminal unsaturated sugar and a product with a
CC  terminal 5'-phosphate.
CC  -1- SUBCELLULAR LOCATION: Nuclear.
CC  -1- TISSUE SPECIFICITY: EXPRESSED IN THE SILIQUES, FLOWERS, AND STEMS.
CC  A HIGH LEVEL EXPRESSION IS SEEN IN THE LEAVES.
CC  -1- SIMILARITY: BELONGS TO THE AP/EXO FAMILY OF DNA REPAIR ENZYMES.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).

```

```

CC  EMBL: AC004625; AAC3731.1;
DR  EMBL: X76912; CAA54234.1;
DR  HSSP: P27695; IE9N.
DR  InterPro: IPR000097; AP_endonuclease_family_1.
DR  InterPro: IPR003034; SAP.
DR  Pfam: PF01260; AP_endonuclease1. 1.
DR  SMART: SM00513; SAP; 1.
DR  PROSITE: PS00726; AP_NUCLEASE_F1_1; 1.
DR  PROSITE: PS00727; AP_NUCLEASE_F1_2; 1.
DR  PROSITE: PS00728; AP_NUCLEASE_F1_3; 1.
FT  DNA repair: lyase; Nuclear protein.
FT  HIGHLY CHARGED; INCREASES THE AFFINITY OF
FT  DOMAIN 1 278
FT  ARP FOR DNA.
FT  DOMAIN 279 536
FT  METAL 313 313 MAGNESIUM OR MANGANESE (BY SIMILARITY).
FT  ACT_SITE 527 527 GENERAL BASE (BY SIMILARITY).
SQ  SEQUENCE 536 AA; 60260 MW; 5C1FC17EA991D27B CRC64;

Query Match      90.5%; Score 38; DB 1; Length 536;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  1 MOPPLP 7
DB  264 MRPPPLP 270

RESULT 4
SERO_GALME STANDARD; PRT; 167 AA.
AC  O76192;
DC  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DE  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Seroin precursor (Silk 23 kDa glycoprotein).
DE  Galleria mellonella (Wax moth).
OS  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC  Pyraloidea; Pyralidae; Galleriinae; Galleria.
OX  NCBI_TaxID=7137;
[1]
RN  SEQUENCE FROM N.A., AND SEQUENCE OF 18-31.
RP  TISSUE=Silk gland;
RC  MEDLINE=98288272; PubMed=9624126;
RA  Zurovec M., Yang C., Kodrik D., Sehnal F.;
RT  *Identification of a novel type of silk protein and regulation of its
RT  expression.
RL  J. Biol. Chem. 273:15423-15428(1998).
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- TISSUE SPECIFICITY: PRODUCED BY BOTH THE POSTERIOR (PSG) AND
CC  MIDDLE (MSG) SECTIONS OF SILK GLANDS.
CC  -1- DEVELOPMENTAL STAGE: SEROIN MRNA IS HIGH IN THE SILK GLANDS OF
CC  COCOON SPINNING, AND THEREAFTER RAPIDLY DROPS TO AN UNDETECTABLE
CC  LEVEL.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: AF009828; AAC25171.1;
KW  Silk; Glycoprotein; Signal; Repeat.
FT  SIGNAL 1 17
FT  CHAIN 18 167 SEROIN.
FT  REPEAT 38 46 1-1.
FT  REPEAT 56 64 1-2.

```

FT REPEAT 76 78 2-1.
 FT REPEAT 79 81 2-2.
 FT REPEAT 82 84 2-3.
 FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 167 AA; 18088 MW; 27A6ABE862774EB9 CRC64;

Query Match 88.1%; Score 37; DB 1; Length 167;
 Best Local Similarity 100.0%; Pred. No. 24;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPEPLP 7
 Db 42 QPEPLP 47

RESULT 5
 ID RGM1_YEAST STANDARD; PRT; 211 AA.
 AC 000453;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable transcription repressor protein RGM1.
 GN RGM1 OR YMR182C OR YMR8010.12C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 ON NCBI_TaxID=4932;
 RX MEDLINE=92020118; PubMed=1923755;
 RA Estruch F.;
 RT "The yeast putative transcriptional repressor RGM1 is a proline-rich
 zinc finger protein.";
 RL Nucleic Acids Res. 19:4873-4877(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- INDUCTION: UNDER THE CONTROL OF THE INDUCIBLE GAL10 PROMOTER.
 CC -1- DOMAIN: THE PRO-RICH REGION OF RGM1 ATTACHED TO A HETEROLOGOUS DNA
 BINDING DOMAIN IS ABLE TO REPRESS THE EXPRESSION OF THE TARGET
 GENE.
 CC -1- SIMILARITY: TO THE MAMMALIAN EGR (EARLY GROWTH RESPONSE) PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL; X59861; CAA42521.1; -
 CC EMBL; Z49808; CAA89915.1; -
 CC PIR; S17249; S17249.
 CC HSSP; P08046; 1A1I.
 CC TRANSFAC; T00730; -
 CC SGD; S0004794; RGM1.
 CC InterPro; IPR000842; Znf-C2H2.
 CC Pfam; PF00096; zf-C2H2; 2.
 CC SMART; SM00355; ZNF_C2H2; 2.
 CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 CC PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
 CC Transcription regulation; Repressor; DNA-binding; Zinc-finger;
 KW Metal-binding; Nuclear protein; Repeat.
 FT DOMAIN 6 11 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT ZN_FING 19 44 C2H2-TYPE.

FT ZN_FING 50 73 C2H2-TYPE.
 FT DOMAIN 95 211 PRO-RICH.
 FT CONFLICT 114 114 A -> V (IN REF. 1).
 SQ SEQUENCE 211 AA; 23855 MW; FADCE4757E7D372 CRC64;

Query Match 88.1%; Score 37; DB 1; Length 211;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPEPLP 7
 Db 154 QPEPLP 159

RESULT 6
 ID FASL_RAT STANDARD; PRT; 278 AA.
 AC P36940;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE FAS antigen ligand.
 GN TNFSF6 OR APTLIG1 OR FASL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RX MEDLINE=94084792; PubMed=7505205;
 RA Suda T., Takahashi T., Golstein P., Nagata S.;
 RT "Molecular cloning and expression of the Fas ligand, a novel member
 of the tumor necrosis factor family.";
 RL Cell 75:1169-1178(1993).
 CC -1- FUNCTION: CYTOKINE THAT BINDS TO FAS ANTIGEN, A RECEPTOR THAT
 TRANSDUCES THE APOPTOTIC SIGNAL INTO CELLS. MAY BE INVOLVED IN
 CYTOTOXIC T CELL MEDIATED APOPTOSIS AND IN T CELL DEVELOPMENT.
 CC FAS-ANTIGEN MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE INDUCTION OF
 PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED SUICIDE OF MATURE
 T CELLS, OR BOTH.
 CC -1- SUBUNIT: HOMODIMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MAY BE RELEASED
 INTO THE EXTRACELLULAR FLUID, PROBABLY BY CLEAVAGE FORM THE CELL
 SURFACE.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ACTIVATED SPLENOCYTES AND
 THYMOCYTES. MODERATE OR WEAK EXPRESSION FOUND IN SMALL INTESTINES,
 KIDNEY AND LUNG.
 CC -1- INDUCTION: BY PMA/IONOMYCIN AND CONCAVALIN/INTERLEUKIN-2.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL; U03470; AAC52129.1; -
 CC HSSP; P01375; 4TSV.
 CC InterPro; IPR003263; TNF_5.
 CC InterPro; IPR003636; TNF_abc.
 CC InterPro; IPR000478; TNF_family.
 CC Pfam; PF00229; TNF; 1.
 CC ProDom; PD002012; TNF_abc; 1.
 CC ProDom; PD008600; TNF_5; 1.
 CC SMART; SM00207; TNF; 1.
 CC PROSITE; PS00251; TNF_1; 1.
 CC PROSITE; PS0049; TNF_2; 1.
 CC Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Apoptosis.
 FT DOMAIN 1 77 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 78 99 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).

FT DOMAIN 100 278 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 4 69 PRO-RICH.
 FT DOMAIN 45 58 POLY-PRO.
 FT DISULFID 199 230 BY SIMILARITY.
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 278 AA; 31140 MW; 2898E18A862CEAC6 CRC64;

Query Match 88.1%; Score 37; DB 1; Length 278;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 111111
 DB 60 OPPPLP 65

RESULT 7
 BLAR_MELGA STANDARD; PRT; 483 AA.
 ID P07700;
 AC 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Beta-1 adrenergic receptor (Beta-1).
 OS Melagris gallopavo (Common turkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
 OX NCBI_TaxID=9103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86313664; PubMed=3018746;
 RA Varden Y., Rodriguez H., Wong S.K.-F., Brandt D.R., May D.C.,
 RA Burnler J., Harkins R.N., Chen E.Y., Ramachandran J., Ullrich A.,
 RA Koss E.N.;
 RT "The avian beta-adrenergic receptor: primary structure and membrane topology."
 RT Proc. Natl. Acad. Sci. U.S.A. 83:6795-6799(1986).
 RN [1]
 RP STRUCTURE BY NMR OF 345-359.
 RX MEDLINE=95129696; PubMed=7828722;
 RA Jung H., Windhaber R., Palm D., Schnackerz K.D.;
 RT "NMR and circular dichroism studies of synthetic peptides derived from the third intracellular loop of the beta-adrenoceptor."
 RT FEBS Lett. 358:133-136(1995).
 CC -1- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-INDUCED ACTIVATION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G PROTEINS. THIS RECEPTOR BINDS EPINEPHRINE AND NOREPINEPHRINE WITH APPROXIMATELY EQUAL AFFINITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- PFM: HOMOLOGOUS DESENSITIZATION OF THE RECEPTOR IS MEDIATED BY ITS PHOSPHORYLATION BY BETA-ADRENERGIC RECEPTOR KINASE.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: M14379; AAA49627.1; -
 CC PIR: A25896; A25896.
 CC PUB; IDEP; 14-OCT-96.
 CC GCRDB; GCR_0164; -
 CC InterPro; IPR000276; GPCR_Rhodopsin.
 CC Pfam; PF00001; 7tm.1; 1.
 CC PRINTS; PR00237; GPCRHOOPS.
 CC PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; 1.
 CC PROSITE; PS50262; G_PROTEIN_RECPT_FL_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Phosphorylation; Lipoprotein; Palmitate;
 KW 3d-structure.
 FT DOMAIN 1 43 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 44 67 1 (POTENTIAL).
 FT TRANSMEM 68 77 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 78 94 2 (POTENTIAL).
 FT TRANSMEM 95 115 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 116 137 3 (POTENTIAL).
 FT TRANSMEM 138 159 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 160 189 4 (POTENTIAL).
 FT DOMAIN 190 205 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 206 228 5 (POTENTIAL).
 FT DOMAIN 229 290 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 291 315 6 (POTENTIAL).
 FT DOMAIN 316 321 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 322 344 7 (POTENTIAL).
 FT DOMAIN 345 483 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 14 14 N-LINKED (GLCNAC. . .) (PROBABLE).
 FT DISULFID 114 192 BY SIMILARITY.
 FT LIPID 358 358 PALMITATE (BY SIMILARITY).
 SQ SEQUENCE 483 AA; 54078 MW; B1A7E71F6CC3E4 CRC64;

Query Match 88.1%; Score 37; DB 1; Length 483;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 111111
 DB 256 OPPPLP 261

RESULT 8
 EXTN_TOBAC STANDARD; PRT; 620 AA.
 ID P13983;
 AC 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Extensin precursor (Cell wall hydroxyproline-rich glycoprotein).
 GN HRP3.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. XANTHI; TISSUE=leaf;
 RA MEDLINE=90128263; PubMed=2612909;
 RX Keller B., Lamb C.J.;
 RT "Specific expression of a novel cell wall hydroxyproline-rich glycoprotein gene in lateral root initiation."
 RT Genes Dev. 3:1639-1646(1989).
 CC -1- FUNCTION: HAS A SPECIALIZED STRUCTURAL FUNCTION, POSSIBLY IN THE MECHANICAL PENETRATION OF THE CORTEX AND EPIDERMIS OF THE MAIN ROOT.
 CC -1- SUBCELLULAR LOCATION: Extracellular matrix.
 CC -1- PFM: EXTENSINS CONTAIN A CHARACTERISTIC REPEAT OF THE PENTAPEPTIDE SER-PRO(4). THE PROLINE RESIDUE IS HYDROXYLATED AND THEN GLYCOSYLATED.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: X13885; CAA32090.1; -
 CC PIR: S06733; S06733.

Db 10 EPPPLP 15

RESULT 29

PGCL_PIG STANDARD; PRT: 193 AA.

AC 095250;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Membrane associated progesterone receptor component 1.

GN PGRMC1 OR PGRMC.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OC NCBI_TaxID=9823;

RP SEQUENCE FROM N.A.

RC TISSUE-Vascular smooth muscle;

RX MEDLINE=97112407; PubMed=8954087;

RA Falkenstein E., Meyer C., Eisen C., Scriba P.C., Wehling M.;

RT "Full-length cDNA sequence of a progesterone membrane-binding protein from porcine vascular smooth muscle cells.";

RL Biochem. Biophys. Res. Commun. 229:86-89(1996).

RN [2]

RP SEQUENCE OF 1-20, AND CHARACTERIZATION.

RC TISSUE-Liver;

RX MEDLINE=96370823; PubMed=8774719;

RA Meyer C., Schmid R., Scriba P.C., Wehling M.;

RT "Purification and partial sequencing of high-affinity progesterone-binding site(s) from porcine liver membranes.";

RL Eur. J. Biochem. 239:726-731(1996).

CC -1- FUNCTION: RECEPTOR FOR PROGESTERONE.

CC -1- SUBUNIT: HETERODIMER OR HETEROTRIMER (POTENTIAL).

CC -1- SUBCELLULAR LOCATION: MICROSOMAL; MEMBRANE-BOUND.

CC -1- SIMILARITY: BELONGS TO THE MAPR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: X99714; CA68050.1; -

DR InterPro: IPR004128; SBD.

DR Pfam: PF03001; SBD: 1.

KW Receptor; Steroid-binding; Transmembrane; Microsome.

FT INIT_MET 0

FT TRANSMEM 24 42

FT SEQUENCE 193 AA; 21478 MW; F7524857C95CDB23 CRC64;

SO

Query Match 81.0%; Score 34; DB 1; Length 193;

Best Local Similarity 83.3%; Pred. No. 76;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 OEPPLP 7

Db 59 EPPPLP 64

RESULT 30

PGCL_HUMAN STANDARD; PRT: 194 AA.

AC 000264; Q9UGJ9;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Membrane associated progesterone receptor component 1 (mPR).

GN PGRMC1 OR PGRMC OR HPR6.6.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OC NCBI_TaxID=9606;

OX (1)

RP SEQUENCE FROM N.A.

RC TISSUE-Liver;

RX MEDLINE=98368653; PubMed=9705155;

RA Gerdes D., Wehling M., Leube B., Falkenstein E.;

RT "Cloning and tissue expression of two putative steroid membrane receptors.";

RL Biol. Chem. 379:907-911(1998).

RN [2]

RP SEQUENCE OF 1-108 FROM N.A.

RA Bernauer S., Wehling M., Falkenstein E.;

RT "The human progesterone membrane binding protein gene: genomic structure and functional analysis.";

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: RECEPTOR FOR PROGESTERONE (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: MICROSOMAL; MEMBRANE-BOUND (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE MAPR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: Y12711; CA73248.1; -

DR EMBL: AJ249131; CAB65109.1; -

DR InterPro: IPR004128; SBD.

DR Pfam: PF03001; SBD: 1.

KW Receptor; Steroid-binding; Transmembrane; Microsome.

FT INIT_MET 0

FT TRANSMEM 24 42

FT SEQUENCE 194 AA; 21540 MW; 558A8AD285FE6AB CRC64;

SO

Query Match 81.0%; Score 34; DB 1; Length 194;

Best Local Similarity 83.3%; Pred. No. 76;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 OEPPLP 7

Db 60 EPPPLP 65

RESULT 31

PGCL_MOUSE STANDARD; PRT: 194 AA.

AC 055022;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Membrane associated progesterone receptor component 1.

GN PGRMC1 OR PGRMC.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.

RC STRAIN-BALB/C; TISSUE-Testis;

RA Kwon S., Lunn R.M., O'Brien D.A., Bell D.A., Eddy E.M.;

RT "The expression of a putative membrane associated progesterone receptor component in the mouse testis and epididymis.";

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: RECEPTOR FOR PROGESTERONE (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: MICROSOMAL; MEMBRANE-BOUND (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE MAPR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AF042491; AAB97466.1; -
 DR MGD; MGI:1858305; Pgrmcl.
 DR InterPro: IPR004128; SBD.
 DR Pfam: PF03001; SBD; 1.
 KW Receptor; Steroid-binding; Transmembrane; Microsome.
 FT INIT_MER 0 0 BY SIMILARITY.
 FT TRANSMEM 24 42 POTENTIAL.
 FT SEQUENCE 194 AA; 21553 MW; 3B14CE24F40ABD40 CRC64;

Query Match 81.0%; Score 34; DB 1; Length 194;
 Best Local Similarity 83.3%; Pred. No. 76;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 Db 60 EPPPLP 65

RESULT 32
 PGCL_RAT STANDARD; PRT; 194 AA.
 AC P70580; 070606;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Membrane associated progesterone receptor component 1 (Acidic 25 kDa
 DE protein) (25-DX).
 GN PGRMC1 OR PGRMC OR LEM1 OR 25DX.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=liver;
 RX MEDLINE=97158736; PubMed=9006096;
 RA Selman O., Lucier G.W., Clark G.C., Tilscher A.M.,
 RA Vanden Heuvel J.P., Gastel J.A., Walker N.J., Suter T.R., Bell D.A.:
 RT Isolation and characterization of a novel gene induced by 2,3,7,8-
 RT tetrachlorodibenzo-p-dioxin in rat liver.*;
 RL Carcinogenesis 17:2609-2615(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FISCHER 344; TISSUE=liver;
 RA Noelle I., Sohn K., Wegehingl S., Wieland F.;
 RT Rat homologue to a putative progesterone binding protein: molecular
 RT characterization and localization.*;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-14.
 RC STRAIN=Wistar; TISSUE=liver;
 RA Hubbard M.J., McHugh N.J.;
 RT "Acidic 25-kDa protein in rat liver microsomes".
 RL Submitted (JUL-1999) to the SWISS-PROT data bank.
 CC -1- FUNCTION: RECEPTOR FOR PROGESTERONE (BY SIMILARITY). MAY BE
 CC IMPLICATED IN TCDD IMMUNOTOXICITY.
 CC -1- SUBCELLULAR LOCATION: MICROSOMAL; MEMBRANE-BOUND (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LUNG, LIVER,
 CC KIDNEY AND BRAIN. LOW IN TESTIS AND SPLEEN. NOT EXPRESSED IN HEART
 CC AND SKELETAL MUSCLE.
 CC -1- INDUCTION: BY DIOXIN.
 CC -1- SIMILARITY: BELONGS TO THE MAPR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U63315; AAB07125.1; -
 DR EMBL: AJ005837; CA06732.1; -
 DR InterPro: IPR004128; SBD.
 DR Pfam: PF03001; SBD; 1.
 KW Receptor; Steroid-binding; Transmembrane; Microsome.
 FT INIT_MER 0 0
 FT TRANSMEM 24 42 POTENTIAL.
 FT CONFLICT 160 166 TERKHHV -> SSPSTITW (IN REF. 1).
 FT CONFLICT 171 173 KEG -> EGA (IN REF. 1).
 FT CONFLICT 177 177 T -> I (IN REF. 1).
 FT CONFLICT 185 194 PDEAARSD -> OKMLGRVTAAGALFLYFAKSFV
 FT SEQUENCE 194 AA; 21467 MW; 87D086A63A6980D CRC64;

Query Match 81.0%; Score 34; DB 1; Length 194;
 Best Local Similarity 83.3%; Pred. No. 76;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 Db 60 EPPPLP 65

RESULT 33
 SODF_NICEPL STANDARD; PRT; 202 AA.
 AC P22302;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Fe], chloroplast (EC 1.15.1.1) (Fragment).
 GN SODB.
 OS Nicotiana glauca (Leadwort-leaved tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV F2;
 RX MEDLINE=91088619; PubMed=2263641;
 RA Van Camp W., Bowler C., Villarroel R., Tsang E.W.T., van Montagu M.,
 RA Inze D.;
 RT "Characterization of iron superoxide dismutase cDNAs from plants
 RT obtained by genetic complementation in Escherichia coli".
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9903-9907(1990).
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -1- COFACTOR: Iron.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- INDUCTION: BY ENVIRONMENTAL CONDITIONS.
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M55909; AAA34074.1; -
 CC PIR: A39267; A39267.

DR HSSP; P09157; 11SC.
 DR Mendel; 290; N1CPL; SODb; 1.
 DR InterPro; IPR001189; SOD_Mt.
 DR Pfam; PF000801; sodfe; 1.
 DR Pfam; PF02777; sodfe.C; 1.
 DR Prodom; P0000475; SOD_Mt; 1.
 DR PROSITE; PS00088; SOD_MN; 1.
 DR Oxidoreductase; Iron; Chloroplast.
 FT NON_TER 1 1
 FT METAL 26 26 IRON (BY SIMILARITY).
 FT METAL 78 78 IRON (BY SIMILARITY).
 FT METAL 162 162 IRON (BY SIMILARITY).
 FT METAL 166 166 IRON (BY SIMILARITY).
 SQ SEQUENCE 202 AA; 23042 MW; 869E79BDD0302C27 CRC64;

Query Match 81.0%; Score 34; DB 1; Length 202;
 Best Local Similarity 71.4%; Pred. No. 79;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPELP 7
 DB 4 LOPPPY 10

RESULT 34
 YP91_CAEEL STANDARD; PRT; 364 AA.
 AC P48458;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative serine/threonine protein phosphatase C06A1.3 in chromosome
 DE II (EC 3.1.3.16).
 GN C06A1.3
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Chordata; Rhabdillida; Rhabditoidea;
 OC Rhabdillidae; Peleodermidae; Caenorhabdillidae.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA McMurtry A.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O - a protein +
 phosphate.
 CC -1- SIMILARITY: BELONGS TO THE PPP FAMILY OF PHOSPHATASES. PP-1
 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; Z49886; CA90052.1; -
 DR HSSP; P08129; 1FJM.
 DR WormRep; C06A1.3; CE02116.
 DR InterPro; IPR000934; Ser_thr_phosphatase.
 DR Pfam; PF00149; Strophosphatase; 1.
 DR PRINTS; PR00114; STPHPRASE.
 DR SMART; SM00156; PP2AC; 1.
 DR PROSITE; PS00125; SER_THR_PHOSPHATASE; 1.
 KW Hypothetical protein; Hydroxase; Iron; Manganese.
 FT METAL 93 93 IRON (BY SIMILARITY).
 FT METAL 95 95 IRON (BY SIMILARITY).
 FT METAL 121 121 IRON AND MANGANESE (BY SIMILARITY).
 FT METAL 153 153 MANGANESE (BY SIMILARITY).
 FT ACT_SITE 154 154 GENERAL ACID (BY SIMILARITY).
 FT METAL 202 202 MANGANESE (BY SIMILARITY).
 FT METAL 277 277 MANGANESE (BY SIMILARITY).

SQ SEQUENCE 364 AA; 41207 MW; 1384EAB0C9729913 CRC64;

Query Match 81.0%; Score 34; DB 1; Length 364;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPELP 7
 DB 325 MOPPELP 331

RESULT 35
 AUP1_MOUSE STANDARD; PRT; 410 AA.
 ID AUP1_MOUSE
 AC P70295;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ancient ubiquitous protein 1 precursor.
 GN AUP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6J;
 RX MEDLINE-96411699; PubMed-8812468;
 RA Jang W., Weber J.S., Bashir R., Bushby K., Meister M.H.;
 RT Aup1, a novel gene on mouse chromosome 6 and human chromosome 2p13.*;
 RL Genomics 36:366-368(1996).
 CC -1- TISSUE SPECIFICITY: UBICUITOUS.
 CC -1- SIMILARITY: BELONGS TO THE AUP1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; U41736; AAC52839.1; -
 DR MGD; MGI:107789; Aup1.
 DR InterPro; IPR003892; CUE.
 DR Pfam; PF02845; CUE; 1.
 DR SMART; SM00546; CUE; 1.
 KW Signal.
 FT SIGNAL 1 37
 FT CHAIN 38 410 POTENTIAL.
 FT CHAIN 38 410 ANCIENT UBICUITOUS PROTEIN 1.
 SQ SEQUENCE 410 AA; 46121 MW; E7D070CEB296BD5B CRC64;

Query Match 81.0%; Score 34; DB 1; Length 410;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPELP 7
 DB 1 MEPPAP 7

RESULT 36
 S3B4_HUMAN STANDARD; PRT; 424 AA.
 ID S3B4_HUMAN
 AC Q15427;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Splicing factor 3b subunit 4 (Spliceosome associated protein 49) (SAP
 DE 49) (SF350) (Pre-mRNA splicing factor SF3b 49 kDa subunit).
 GN SF3B4 OR SAP49.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NC NCB1_TaxID=9606;
 RN (1)
 RP MEDLINE-95047348; PubMed-7958871;
 RX MEDLINE-95047348; PubMed-7958871;
 RA Champion-Arnaud P., Reed R.;
 RT "The pre-spliceosome components SAP 49 and SAP 145 interact in a
 RT complex implicated in tethering U2 snRNP to the branch site.";
 RT Genes Dev. 8:1974-1983(1994).
 RN (2)
 RP CHARACTERIZATION OF THE SPLICEOSOME.
 RX MEDLINE-20337962; PubMed-10882114;
 RA Das R., Zhou Z., Reed R.;
 RT "Functional association of U2 snRNP with the ATP-independent
 RT spliceosomal complex E.";
 RT Mol. Cell 5:779-787(2000).
 CC -1- FUNCTION: SUBUNIT OF THE SPLICING FACTOR SF3B REQUIRED FOR 'A'
 CC COMPLEX ASSEMBLY FORMED BY THE STABLE BINDING OF U2 snRNP TO THE
 CC BRANCHPOINT SEQUENCE (BPS) IN PRE-MRNA. SEQUENCE INDEPENDENT
 CC BINDING OF SF3A/SF3B COMPLEX UPSTREAM OF THE BRANCH SITE IS
 CC ESSENTIAL. IT MAY ANCHOR U2 snRNP TO THE PRE-MRNA. MAY ALSO BE
 CC INVOLVED IN THE ASSEMBLY OF THE 'E' COMPLEX. SF3B4 HAS BEEN
 CC FOUND IN COMPLEX 'B' AND 'C' AS WELL. BELONGS ALSO TO THE
 CC MINOR U12-DEPENDENT SPLICEOSOME, WHICH IS INVOLVED IN THE SPLICING
 CC OF RARE CLASS OF NUCLEAR PRE-MRNA INTRON.
 CC -1- SUBUNIT: COMPONENT OF SPLICING FACTOR SF3B WHICH IS COMPOSED OF
 CC FOUR SUBUNITS: SF3B4/SAP49, SF3B3/SAP130, SF3B2/SAP145,
 CC SF3B1/SAP155. SF3B ASSOCIATES WITH THE SPLICING FACTOR SF3A AND A
 CC 12S RNA UNIT TO FORM THE U2 snRNP. NUCLEAR RIBONUCLEOPROTEINS
 CC COMPLEX (U2 snRNP). SF3B4 INTERACTS DIRECTLY WITH SF3B2.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: STRONG, TO C.ELEGANS C08B11.5 AND, PARTIAL, TO
 CC YEAST HSH49.
 CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
 CC
 CC "This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L35013; AAA60300.1;
 DR HSP: P19338; 2SXL.
 DR MIM: 605593;
 DR InterPro: IPR00504; RRM.
 DR Pfam: PF00076; rrm; 2.
 DR SMART: SMO0360; RRM; 2.
 DR PROSITE: PSS0102; RRM; 2.
 DR PROSITE: PSS0030; RRM_RNP_1; 1.
 KW Spliceosome; mRNA processing; mRNA splicing; Nuclear protein;
 KW RNA-binding; Repeat.
 FT DOMAIN 13
 FT DOMAIN 13
 FT DOMAIN 100 179 RNA-BINDING (RRM) 1.
 FT DOMAIN 215 218 RNA-BINDING (RRM) 2.
 FT DOMAIN 262 268 POLY-PRO.
 FT DOMAIN 262 268 POLY-PRO.
 SQ SEQUENCE 424 AA; 44385 MW; 212472A25D3FF002 CRC64;

Query Match 81.0%; Score 34; DB 1; Length 424;
 Best Local Similarity 71.4%; Pred. No. 1.7e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MPPPLP 7
 Db 261 MPPPLP 267

RESULT 37
 YU87_MYCTU

ID YU87_MYCTU STANDARD; PRT; 472 AA.
 AC 053304;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 52.6 kDa protein RV3087.
 GN RV3087 OR MT3172 OR MT013.08.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
 NC NCB1_TaxID=1773;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE-98295987; PubMed-9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skellon S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
 RT "Deciphering the Biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RT Nature 393:537-544(1998).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonyay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gali J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RT Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 RL
 CC -1- SIMILARITY: BELONGS TO THE UPF0089 FAMILY.
 CC
 CC "This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AL021309; CA16145.1;
 DR EMBL: AE007134; AAK47508.1; ALT_INIT.
 DR TIGR: MT3172;
 DR Tuberculist; RV3087;
 DR InterPro: IPR004255; UPF0089.
 DR Pfam: PF03007; UPF0089; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 472 AA; 52597 MW; AC03BDB4970FCC0 CRC64;

Query Match 81.0%; Score 34; DB 1; Length 472;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 QPPPLP 7
 Db 170 EPPPLP 175

RESULT 38
 ID SOX9_PIG STANDARD; PRT; 509 AA.
 AC 018896;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)

```

DE Transcription factor SOX-9.
GN SOX9.
OS Sus scrofa (Pig).
OC Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE FROM N.A.
RA Pilon N., Silversides D.W.;
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE NORMAL SKELETAL
CC DEVELOPMENT. MAY REGULATE THE EXPRESSION OF OTHER GENES INVOLVED
CC IN CHONDROGENESIS BY ACTING AS A TRANSCRIPTION FACTOR FOR THESE
CC GENES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: CONTAINS 1 HMG BOX.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL; AF029696; AAB1431.1; -.
DR HSSP; P48436; 1SX9.
DR InterPro; IPR000910; HMG_12_box.
DR Pfam; PF00505; HMG_box; 1.
DR SMART; SM00398; HMG; 1.
KM Transcription regulation; DNA-binding; Nuclear protein.
FT DNAStrand 105 173
FT DOMAIN 339 378 GLN/PRO-RICH.
FT DNAStrand 342 349 POLY-PRO.
SQ SEQUENCE 509 AA; 56267 MW; 069F77A0510BBFC2 CRC64;

Query Match 81.0%; Score 34; DB 1; Length 509;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPVP 7
DB 351 OPPPVP 356

RESULT 39
DABI_HUMAN STANDARD; PRT; 555 AA.
ID DABI_HUMAN
AC 075553; Q9MYA8;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Disabled homolog 1.
GN DABI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-99009343; PubMed-9790777;
RL Lambert de Rouvroit C., Goffinet A.M.;
RT Cloning of human DABI and mapping to chromosome 1p31-p32.;
RL Genomics 53:246-247(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Fazili Z., Sun W., Xu X.-X.;
RT "Aberrant disabled-1 expression in tumors.";
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Adaptor molecule functioning in neural development (By
CC similarity).
CC -1- SUBUNIT: Associates with the SH2 domains of Src, Fyn and Abl (By

```

```

CC similarity).
CC -1- DOMAIN: THE PID DOMAIN SPECIFICALLY BINDS TO THE ASN-PRO-XAA-
CC TYR(P) MOTIF FOUND IN MANY TYROSINE-PHOSPHORYLATED PROTEINS.
CC -1- PTM: Phosphorylated on tyrosine (By similarity).
CC -1- SIMILARITY: CONTAINS 1 PID DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL; AF071062; AAC70068.1; -.
DR EMBL; AF263547; AAF73058.1; -.
DR MIM; 603448; -.
DR InterPro; IPR000050; PID_domain.
DR Pfam; PF00640; PID; 1.
DR SMART; SM00462; PTB; 1.
DR PROSITE; PS01179; PID; 1.
KM Developmental protein; Neurogenesis; Phosphorylation.
FT DOMAIN 36 189
FT CONFLICT 86 86 F -> L (IN REF. 2).
FT CONFLICT 240 241 MISSING (IN REF. 2).
FT CONFLICT 291 291 L -> F (IN REF. 2).
FT CONFLICT 298 298 P -> S (IN REF. 2).
FT CONFLICT 326 326 H -> Q (IN REF. 2).
FT CONFLICT 335 335 G -> R (IN REF. 2).
FT CONFLICT 379 379 M -> V (IN REF. 2).
FT CONFLICT 553 553 D -> A (IN REF. 2).
SQ SEQUENCE 555 AA; 59979 MW; A4B133A0C3AF2034 CRC64;

Query Match 81.0%; Score 34; DB 1; Length 555;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPVP 7
DB 422 OPPPVP 427

RESULT 40
DABI_MACFA STANDARD; PRT; 555 AA.
ID DABI_MACFA
AC 09BGX5;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Disabled homolog 1.
GN DABI.
OS Macaca fascicularis (Cray eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-Frontal cortex;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Teruo K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Adaptor molecule functioning in neural development (By
CC similarity).
CC -1- SUBUNIT: Associates with the SH2 domains of Src, Fyn and Abl (By
CC similarity).
CC -1- DOMAIN: THE PID DOMAIN SPECIFICALLY BINDS TO THE ASN-PRO-XAA-
CC TYR(P) MOTIF FOUND IN MANY TYROSINE-PHOSPHORYLATED PROTEINS.
CC -1- PTM: Phosphorylated on tyrosine (By similarity).
CC -1- SIMILARITY: CONTAINS 1 PID DOMAIN.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB055282; BAB21906.1; -
CC InterPro: IPR000050; PID_domain.
CC Pfam: PF00640; PID: 1.
CC SMART: SM00462; PTB: 1.
CC PROSITE: PS01179; PID: 1.
CC Developmental protein; Neurogenesis; Phosphorylation.
CC KW DOMAIN 36 189 PID.
CC SQ SEQUENCE 555 AA; 59920 MW; A3802A1D0C4EE51 CRC64;

Query Match
Best Local Similarity 81.0%; Score 34; DB 1; Length 555;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
Db 422 QPPVP 427

RESULT 41
FXOL_MOUSE STANDARD; PRT; 652 AA.
AC 09R1E0:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Forkhead box protein O1A (Forkhead in rhabdomyosarcoma).
GN FOXO1A OR FOXO1 OR FKHR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId=10090;
OC [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99278356; PubMed-10347145;
RA Nakae J., Park B.C., Accilli D.;
RT Insulin stimulates phosphorylation of the forkhead transcription
RT factor FKHR on serine 253 through a Wortmannin-sensitive pathway.*;
RT J. Biol. Chem. 274:15982-15985(1999).
CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- PTM: PHOSPHORYLATED BY AKT; INSULIN-INDUCED.
CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF126056; AAD40636.1; -
CC MGD: MGI:1690077; Foxo1.
CC InterPro: IPR001766; Fork_head.
CC Pfam: PF00250; Fork_head_1.
CC PRINTS: PR00053; FORKHEAD.
CC SMART: SM00339; FH: 1.
CC PROSITE: PS00657; FORK_HEAD_1; FALSE_NEG.
CC PROSITE: PS00658; FORK_HEAD_2; 1.
CC PROSITE: PS00039; FORK_HEAD_3; 1.
CC Transcription regulation; DNA-binding; Nuclear protein;
CC Phosphorylation.
CC KW DOMAIN 89 96 POLY-ALA.
CC FT

```

```

FT DOMAIN 135 139 POLY-ALA.
FT DNA_BIND 156 232 FORK-HEAD.
FT MOD_RES 253 253 PHOSPHORYLATION (BY AKT).
SQ SEQUENCE 652 AA; 69502 MW; 3FE4C322AA85205F CRC64;

Query Match
Best Local Similarity 81.0%; Score 34; DB 1; Length 652;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
Db 127 QPPVP 132

RESULT 42
AT12_HSV1 STANDARD; PRT; 718 AA.
ID AT12_HSV1
AC P10230:
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Alpha trans-inducing factor 78 kDa protein.
GN UL46.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OC NCBI_TaxId=10299;
OC [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88274327; PubMed-2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1.*;
RT J. Gen. Virol. 69:1531-1574(1988).
CC -1- FUNCTION: MODULATE ALPHA TRANS-INDUCING FACTOR-DEPENDENT
CC ACTIVATION OF ALPHA GENES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10879; BAA01692.1; -
CC DR PIR: A30089; TNBER6.
CC Transcription regulation; Trans-acting factor.
CC KW SEQUENCE 718 AA; 78244 MW; 4E376A2A74474FBE CRC64;
CC SQ

Query Match
Best Local Similarity 81.0%; Score 34; DB 1; Length 718;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
Db 467 EPPPLP 472

RESULT 43
EBN6_EBV STANDARD; PRT; 992 AA.
ID EBN6_EBV
AC P03204:
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE EBNA-6 nuclear protein (EBNA-3C) (EBNA-4B).
GN BBRF3-BERF4.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC FT

```

CC Gammaherpesvirinae; Lymphocryptovirus.
 OX NCBI_TaxID=10377;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84270667; PubMed=6087149;
 RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.T.,
 RA Gibson T.J., Hatfull G., Hudson G.S., Sachwell S.C., Seguin C.,
 RA Tuffnell P.S., Barrett B.G.;
 RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
 RL Nature 310:207-211(1984).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=88155772; PubMed=2831394;
 RA Pettit L., Sample J., Wang F., Kieff E.;
 RT "A fifth Epstein-Barr virus nuclear protein (EBNA5C) is expressed in
 RT latently infected growth-transformed lymphocytes.";
 RL J. Virol. 62:1330-1338(1988).
 RN [3]
 RP SUBCELLULAR LOCATION.
 RX MEDLINE=90266473; PubMed=2161150;
 RA Pettit L., Sample C., Kieff E.;
 RT "Subnuclear localization and phosphorylation of Epstein-Barr virus
 RT latent infection nuclear proteins.";
 RL Virology 176:563-574(1990).
 CC -1- FUNCTION: INVOLVED IN LATENT CYCLE.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
 CC -1- SIMILARITY: SOME SIMILARITIES EXIST BETWEEN EBNA 4, 5, AND 6.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 CC EMBL: V01555; CAA24859.1;
 DR PIR: A03766; Q08E25.
 KW Nuclear protein; Repeat.
 FT DOMAIN 74 80 POLY-ARG.
 FT DOMAIN 551 610 10 X 5 AA TANDEM REPEATS.
 FT DOMAIN 741 779 3 X 13 AA TANDEM REPEATS.
 FT SEQUENCE 992 AA; 109129 MW; 39BEAB9BC515BD84 CRC64;
 SQ
 Query Match 81.0%; Score 34; DB 1; Length 992;
 Best Local Similarity 83.3%; Pred. No. 4.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 2 QPEPLP 7
 DB 479 QPEPLP 484
 RESULT 44
 PTC2_MOUSE STANDARD; PRT: 1182 AA.
 ID PTC2_MOUSE
 AC 035595;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Patched protein homolog 2 (PTC2).
 GN PTC2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Embryo;
 RX MEDLINE=98122566; PubMed=9462734;
 RA Motoyama J., Takabatake T., Takeshima K., Hui C.-C.;
 RT "Ptc2, a second mouse patched gene is co-expressed with Sonic

RT hedgehog.";
 RL Nat. Genet. 18:104-106(1998).
 RN [2]
 RP SEQUENCE OF 196-446 FROM N.A.
 RC STRAIN-BALB/C; TISSUE-Neuroretina;
 RX MEDLINE=97379366; PubMed=9237688;
 RA Takabatake T., Ogawa M., Takahashi T.C., Mizuno M., Okamoto M.,
 RA Takeshima K.;
 RT "Hedgehog and patched gene expression in adult ocular tissues.";
 RL FEBS Lett. 410:485-489(1997).
 CC -1- FUNCTION: MAY HAVE A ROLE IN EPIDERMAL DEVELOPMENT. MAY ACT AS A
 CC RECEPTOR FOR SONIC HEDGEHOG (SHH).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL CELLS OF THE
 CC DEVELOPING HAIR, TOOTH AND WHISKER.
 CC -1- DEVELOPMENTAL STAGE: DETECTED IN 8.5 TO 17.5 DAYS POST COITUM
 CC EMBRYOS.
 CC -1- SIMILARITY: BELONGS TO THE PATCHED FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 CC EMBL: AB010833; BAA24691.1;
 DR MGI: 1095405; Ptc2.
 DR InterPro: IPR000731; HMGR-patched_5TM.
 DR InterPro: IPR003392; Patched.
 DR Pfam: PF02460; Patched; 1.
 DR PROSITE: PS50156; SSD: 1.
 KW Receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 58 78 POTENTIAL.
 FT DOMAIN 79 394 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 395 414 POTENTIAL.
 FT DOMAIN 415 428 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 429 449 POTENTIAL.
 FT DOMAIN 450 457 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 458 478 POTENTIAL.
 FT DOMAIN 479 501 POTENTIAL.
 FT TRANSMEM 502 522 POTENTIAL.
 FT DOMAIN 523 531 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 532 552 POTENTIAL.
 FT DOMAIN 553 686 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 687 707 POTENTIAL.
 FT DOMAIN 708 963 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 964 984 POTENTIAL.
 FT DOMAIN 985 991 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 992 1012 POTENTIAL.
 FT DOMAIN 1013 1013 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1014 1034 POTENTIAL.
 FT DOMAIN 1035 1064 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 1065 1085 POTENTIAL.
 FT DOMAIN 1086 1086 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1087 1107 POTENTIAL.
 FT DOMAIN 1108 1182 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 1182 AA; 128585 MW; 715233D912C352F2 CRC64;
 SQ
 Query Match 81.0%; Score 34; DB 1; Length 1182;
 Best Local Similarity 71.4%; Pred. No. 4.9e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 MOPPLP 7
 DB 1164 LHPPPLP 1170

```

RESULT 45
PER2_HUMAN STANDARD: PRT: 1255 AA.
ID PER2_HUMAN
AC O15055;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Period circadian protein 2.
GN PER2 OR KIAA0347.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka K.-I., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
RN [2]
RP REVISIONS TO C-TERMINUS.
RC TISSUE=Brain;
RX Nagase T., Ishikawa K.-I., Seki N., Nakajima D., Ohira M.,
RA Miyajima N., Kotani H., Nomura N., Ohara O.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBD databases.
RN [3]
RP TISSUE EXPRESSION.
RX MEDLINE=98087121; PubMed=9427249;
RA Shearman L.P., Zylka M.J., Weaver D.R., Kolakowski L.F. Jr.,
RA Reppert S.M.;
RT "Two period homologs: circadian expression and photic regulation in
RT the suprachiasmatic nuclei.";
RL Neuron 19:1261-1269(1997).
CC -1- FUNCTION: CIRCADIAN REGULATOR THAT MAY ACT AS A TRANSCRIPTION
CC FACTOR. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN TRANSCRIPTIONAL
CC LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING INDIRECT
CC TRANSCRIPTIONAL INHIBITION. EXPRESSION OSCILLATES WITH A 24 HR
CC RHYTHM IN THE SUPRACHIASMATIC NUCLEUS (SCN) AND THE WHOLE EYES.
CC OSCILLATIONS ARE MAINTAINED UNDER CONSTANT DARKNESS AND ARE
CC RESPONSIVE TO CHANGES OF THE LIGHT/DARK CYCLES. THERE IS A 4 HOUR
CC TIME DELAY BETWEEN PER1 AND PER2 OSCILLATIONS. THE EXPRESSION
CC RHYTHMS APPEAR TO ORIGINATE FROM RETINA (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. FOUND IN HEART, BRAIN,
CC PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY AND PANCREAS.
CC -1- INDUCTION: BY LIGHT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AB002345; BAA20804.2; ALT_INT.
CC MIM: 603426;
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000014; PAC.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 2.
DR PROSITE: PS00112; PAS; 1.
KM Transcription regulation. Nuclear protein. Repeat. Biological rhythms.
FT DOMAIN 109 146 Helix-loop-helix motif (by similarity).
FT DOMAIN 182 248 PAS 1.

```

```

FT DOMAIN 319 389 PAS 2.
FT DOMAIN 398 438 PAC.
FT DOMAIN 510 513 POLY-ARG.
FT DOMAIN 789 806 NUCLEAR LOCALIZATION SIGNAL.
FT DOMAIN 842 979 PRO-RICH.
SQ SEQUENCE 1255 AA; 136579 MW; 2AEF2C6BD4B6CB0 CRC64;

Query Match
Best Local Similarity 81.0%; Score 34; DB 1; Length 1255;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPPPLP 7
DB 891 VQPPPP 897

RESULT 46
YBE7_YEAST STANDARD: PRT: 1381 AA.
ID YBE7_YEAST
AC P34216;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 150.8 kDa protein in SEC17-OCRI intergenic region.
GN YBL047C OR YBL0520 OR YBL0501.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE OF 1-961 FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=95176707; PubMed=7871888;
RA de Wergifosse P., Jacques B., Jomiaux J.-L., Purnelle B., Skala J.,
RA Goffeau A.;
RT "The sequence of a 22.4 kb DNA fragment from the left arm of yeast
RT chromosome II reveals homologues to bacterial proline synthetase and
RT murine alpha-adaptin, as well as a new permease and a DNA-binding
RT protein.";
RL Yeast 10:1489-1496(1994).
RN [2]
RP SEQUENCE OF 579-1381 FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=94205266; PubMed=8154187;
RA Scherens B., el Bakoury M., Vierendeels F., Dubois E., Messenguy F.;
RT "Sequencing and functional analysis of a 32,560 bp segment on the
RT left arm of yeast chromosome II. Identification of 26 open reading
RT frames, including the KIP1 and SEC17 genes.";
RL Yeast 9:1355-1371(1993).
CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: Z35808; CAA84867.1;
CC EMBL: X78214; CAA55048.1;
CC EMBL: Z23261; CAA80797.1;
CC PIR: S45781; S45781.
CC SGD: S0000143; YBL047C.
DR InterPro: IPR000261; EPS15_repeat.
DR InterPro: IPR000449; UBA.
DR Pfam: PF00036; ethand; 2.
DR Pfam: PF00627; UBA; 1.
DR SMART: SM00054; EPH; 2.
DR SMART: SM00027; EH; 3.
DR SMART: SM00165; UBA; 1.
KM Hypothetical protein.

```

SO SEQUENCE 1381 AA; 150783 MW; 626FD261DC8A7D99 CRC64;

Query Match 81.0%; Score 34; DB 1; Length 1381;
Best Local Similarity 83.3%; Pred. No. 5.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
DB 517 QPPPLP 522

RESULT 47
ZAP3_MOUSE STANDARD; PRT; 1386 AA.
AC 09R017;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nucleolar protein ZAP3.
GN ZAP3 OR ZAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal Liver;
RA Misawa K., Nosaka T., Kitamura T.;
RT "A huge nuclear protein rich in proline similar to human hypothetical
RT protein zap3 and zap13.";
RT Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB031168; BAA85182.1; -
DR MGD; MGT1926195; Zap3.
KW Nucleolar protein.
FT DOMAIN 15 204 PRO-RICH.
FT DOMAIN 355 473 GLN-RICH.
FT DOMAIN 925 1012 ARG-RICH.
SQ SEQUENCE 1386 AA; 155130 MW; D862F9918ED21DF CRC64;

Query Match 81.0%; Score 34; DB 1; Length 1386;
Best Local Similarity 83.3%; Pred. No. 5.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
DB 1043 QPPPLP 1048

RESULT 48
PK3B_HUMAN STANDARD; PRT; 1634 AA.
AC 000750; O95666;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phosphatidylinositol 3-kinase C2 domain-containing beta polypeptide
DE (EC 2.7.1.137) (Phosphoinositide 3-kinase-C2-beta) (PtdIns-3-kinase
DE C2 beta) (PI3K-C2beta) (C2-PI3K).
GN PI3KCB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Breast;
RX MEDLINE=97289668; PubMed=9144573;
RA Brown R.A., Ho L.K.F., Weber-Hall S.J., Shipley J.M., Fry M.J.;
RT "Identification and cDNA cloning of a novel mammalian C2 domain-
RT containing phosphoinositide 3-kinase, Hsc2-PI3K.";
RT Biochem. Biophys. Res. Commun. 233:537-544(1997).
RN [2]
RP SEQUENCE OF 26-1634 FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Monocytes;
RX MEDLINE=99047700; PubMed=9830063;
RA Arcaro A., Volinia S., Zvelebil M.J., Stein R., Watton S.J.,
RA Layton M.J., Gout I., Ahmed K., Downard J., Waterfield M.D.;
RT "Human phosphoinositide 3-kinase C2beta, the role of calcium and the
RT C2 domain in enzyme activity.";
RL J. Biol. Chem. 273:33082-33090(1998).
CC -1- FUNCTION: PHOSPHORYLATES PTDINS AND PTDINS4P WITH A PREFERENCE FOR
CC PTDINS. DOES NOT PHOSPHORYLATE PTDINS(4,5)P2.
CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol -> ADP +
CC 1-phosphatidyl-1D-myo-inositol 3-phosphate.
CC -1- SUBCELLULAR LOCATION: FOUND MOSTLY IN THE MICROsome, BUT ALSO IN
CC THE PLASMA MEMBRANE AND CYTOSOL.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED, BUT LEVELS ARE HIGHEST IN
CC THYMUS AND PLACENTA AND LOWEST IN PERIPHERAL BLOOD, SKELETAL
CC MUSCLE AND KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-26 IS THE INITIATOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y11312; CAA72168.1; -
DR EMBL: Y13892; CAA74194.1; -
DR HSSP: P21707; IRSY.
DR MTM: 602838; -
DR InterPro: IPR000008; C2.
DR InterPro: IPR002420; PI3K_C2.
DR InterPro: IPR000341; PI3K_ras_bind.
DR InterPro: IPR001263; PI3K.
DR InterPro: IPR000403; PI3_P14_kinase.
DR InterPro: IPR001683; PX.
DR Pfam: PF00168; C2; 1.
DR Pfam: PF00613; PI3Ka; 1.
DR Pfam: PF00792; PI3K_C2; 1.
DR Pfam: PF00794; PI3K_rbd; 1.
DR Pfam: PF00454; PI3_P14_kinase; 1.
DR Pfam: PF00787; PX; 1.
DR SMART: SM00239; C2; 1.
DR SMART: SM00142; PI3K_C2; 1.
DR SMART: SM00144; PI3K_rbd; 1.
DR SMART: SM00145; PI3Ka; 1.
DR SMART: SM00146; PI3KC; 1.
DR SMART: SM00312; PX; 1.
DR PROSITE: PS00499; C2_DOMAIN_1; FALSE_NEG.
DR PROSITE: PS50004; C2_DOMAIN_2; 1.
DR PROSITE: PS00915; PI3_4_KINASE_1; 1.
DR PROSITE: PS00916; PI3_4_KINASE_2; 1.
DR PROSITE: PS50290; PI3_4_KINASE_3; 1.
KW Transferase; Kinase; Multigene family.
FT DOMAIN 156 162 PRO-RICH.
FT DOMAIN 169 174 PRO-RICH.
FT DOMAIN 1079 1343 PI3K/PI4K.
FT DOMAIN 1517 1608 C2 DOMAIN.
FT CONFLICT 63 63 P -> S (IN REF. 2).

FT CONFLICT 75 75 R -> W (IN REF. 2).
 FT CONFLICT 99 99 Q -> L (IN REF. 2).
 FT CONFLICT 246 246 V -> A (IN REF. 2).
 FT CONFLICT 278 278 K -> E (IN REF. 2).
 FT CONFLICT 567 567 P -> S (IN REF. 2).
 FT CONFLICT 664 665 DM -> EL (IN REF. 2).
 FT CONFLICT 1634 AA; 184856 MW; COB5DF63C668B824 CRC64;
 SO SEQUENCE

Query Match
 Best Local Similarity 81.0%; Score 34; DB 1; Length 1634;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 : |||||
 Db 154 LSPPLP 160

RESULT 49
 ATFS_HUMAN STANDARD; PRT; 282 AA.
 ID ATFS_HUMAN
 AC 09Y2D1; 09UNO3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Cyclic AMP-dependent transcription factor ATF-5 (Activating
 transcription factor 5) (Transcription factor ATF5).
 GN ATF5 OR ATF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RX MEDLINE-20558615; PubMed-11087824;
 RA White J.H., McIlhinney R.A.J., Wise A., Ciruela F., Chan W.Y.,
 RA Emson P.C., Billington A., Marshall F.H.,
 RT "The GABAB receptor interacts directly with the related transcription
 factors CREB2 and ATF5".
 RL Proc. Natl. Acad. Sci. U.S.A. 97:13967-13972(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kohroki J., Tanaka K.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 161-282 FROM N.A.
 RX MEDLINE-99303793; PubMed-10373550;
 RA Patl D., Weistrich M.L., Pion S.E.;
 RT "Human Cdc34 and Rad6 ubiquitin-conjugating enzymes target repressors
 of cyclic AMP-induced transcription for proteolysis".
 RL Mol. Cell. Biol. 19:5001-5013(1999).
 CC -1- FUNCTION: THIS PROTEIN BINDS THE CAMP RESPONSE ELEMENT (CRE)
 (CONSENSUS: 5'-GTGACGT(A/C)(A/G)-3'), A SEQUENCE PRESENT IN MANY
 VIRAL AND CELLULAR PROMOTERS.
 CC -1- SUBUNIT: BINDS DNA AS A DIMER (POTENTIAL). INTERACTS WITH GABAB
 RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: TO OTHER BZIP PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: AB021663; AAC22558.1;
 DR EMBL: AB021663; BAA78477.2;
 DR EMBL: AF101388; AAD28370.1;
 DR MIM: 606398;
 DR InterPro: IPR002637; Hamjp_1like.
 DR InterPro: IPR001871; bzip.

DR pfam: PF00170; bzip; 1.
 DR ProDom: PD004952; Hamjp_1like; 1.
 DR SMART: SM00338; BRU2; 1.
 DR PROSITE: PS00036; BZIP_BASIC; 1.
 KM Transcription regulation; DNA-binding; Activator; Nuclear protein;
 KM Multigene family.
 FT DOMAIN 123 139
 FT DOMAIN 186 194
 FT DNA_BIND 210 230
 FT DOMAIN 236 250
 FT CONFLICT 161 163
 FT SEQUENCE 282 AA; 30674 MW; DDB2F907CA0215A0 CRC64;

Query Match
 Best Local Similarity 78.6%; Score 33; DB 1; Length 282;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 : |||||
 Db 130 LPPPLP 136

RESULT 50
 DHPI_HUMAN STANDARD; PRT; 299 AA.
 ID DHPI_HUMAN
 AC 092874;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNase I homolog protein DHPI precursor (EC 3.1.21.-).
 GN DNASE I OR DNASE I2 OR DHPI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RX MEDLINE-97349121; PubMed-9205125;
 RA Rodriguez A.M., Rodin D., Nomura H., Morton C.C., Weremowicz S.,
 RA Schneider M.C.;
 RT "Identification, localization, and expression of two novel human
 genes similar to deoxyribonuclease I".
 RL Genomics 42:507-513(1997).
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE DNASE I FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: U62647; AAB63981.1;
 DR HSP: P00639; 2DNT.
 DR MIM: 602622;
 DR InterPro: IPR001582; DNase_I.
 DR pfam: PF01181; DNase_I; 1.
 DR PRINTS: PR00130; DNASEI.
 DR ProDom: PD005408; DNase_I; 1.
 DR SMART: SM00476; DNaseIc; 1.
 DR PROSITE: PS00918; DNASE_I_2; 1.
 DR PROSITE: PS00919; DNASE_I_1; 1.
 KM Hydrolase; Endonuclease; Nuclease; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 299
 FT DISULFID 209 245
 FT ACT_SITE 99 99
 FT ACT_SITE 170 170
 FT SEQUENCE 299 AA; 32853 MW; 1B66A5590D33D0A6 CRC64;

Query Match 78.6%; Score 33; DB 1; Length 299;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MOPPP 7
: |||||
DB 152 LMPPLP 158

Search completed: August 8, 2002, 07:49:28
Job time: 121 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2002, 07:46:47 ; Search time 24.97 Seconds
(without alignments)
48.497 Million cell updates/sec

Title: US-09-641-801-1
Perfect score: 42
Sequence: 1 MQPPLP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 150 summaries

Database :

SPREMBL_19.*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriophage:*
17: sp.archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	316	5	09VAX5
2	40	95.2	1776	5	09MA20
3	39	92.9	612	16	09PF38
4	39	92.9	1404	5	09GRR6
5	39	92.9	1444	11	09JLPO
6	38	90.5	542	10	09SGY5
7	37	88.1	72	15	087610
8	37	88.1	73	15	087603
9	37	88.1	99	15	087482
10	37	88.1	126	5	09NDS3
11	37	88.1	126	5	09NDS2
12	37	88.1	164	5	018043
13	37	88.1	178	10	048681
14	37	88.1	186	5	020796
15	37	88.1	294	10	09L653
16	37	88.1	322	4	09BSM2

17	37	88.1	346	11	099LN2	099LN2 mus musculu
18	37	88.1	357	4	096R24	096R24 homo sapien
19	37	88.1	393	10	022921	022921 arabidopsis
20	37	88.1	421	10	09FTS7	09FTS7 oryza sativ
21	37	88.1	440	5	096M1	096M1 drosophila
22	37	88.1	456	5	09N06	09N06 leishmania
23	37	88.1	471	5	093571	093571 caenorhabd
24	37	88.1	588	10	022717	022717 arabidopsis
25	37	88.1	629	11	099NA4	099NA4 mus musculu
26	37	88.1	632	10	093254	093254 arabidopsis
27	37	88.1	651	4	09R25	09R25 homo sapien
28	37	88.1	667	11	035745	035745 mus musculu
29	37	88.1	675	11	09EP73	09EP73 mus musculu
30	37	88.1	692	5	044760	044760 caenorhabd
31	37	88.1	704	11	09JLQ4	09JLQ4 mus musculu
32	37	88.1	721	10	094C37	094C37 arabidopsis
33	37	88.1	728	11	09ES78	09ES78 mus musculu
34	37	88.1	731	4	096S74	096S74 homo sapien
35	37	88.1	733	11	0922Y1	0922Y1 mus musculu
36	37	88.1	738	11	091Z55	091Z55 mus musculu
37	37	88.1	749	10	023045	023045 arabidopsis
38	37	88.1	750	4	09BRR3	09BRR3 homo sapien
39	37	88.1	819	5	09VHR3	09VHR3 drosophila
40	37	88.1	830	5	093176	093176 caenorhabd
41	37	88.1	838	5	090919	090919 caenorhabd
42	37	88.1	838	5	0950Y9	0950Y9 caenorhabd
43	37	88.1	926	5	09NFD4	09NFD4 drosophila
44	37	88.1	961	5	046037	046037 drosophila
45	37	88.1	962	5	024584	024584 drosophila
46	37	88.1	1125	5	093203	093203 caenorhabd
47	37	88.1	1154	4	013088	013088 homo sapien
48	37	88.1	1179	11	09JK71	09JK71 rat mus norv
49	37	88.1	1263	11	09QY22	09QY22 mus musculu
50	37	88.1	1291	4	09GFV0	09GFV0 homo sapien
51	37	88.1	1359	4	09UPR1	09UPR1 homo sapien
52	37	88.1	1364	4	075092	075092 homo sapien
53	37	88.1	1475	5	09VE13	09VE13 drosophila
54	37	88.1	1560	5	060275	060275 homo sapien
55	37	88.1	1629	5	09VDM3	09VDM3 drosophila
56	37	88.1	1654	4	09P1Y6	09P1Y6 homo sapien
57	37	88.1	1669	5	09BR99	09BR99 drosophila
58	37	88.1	1912	12	09J5C3	09J5C3 fowlpox vir
59	37	88.1	3312	4	09NT07	09NT07 homo sapien
60	36	85.7	235	10	09EY74	09EY74 zea mays (m
61	36	85.7	318	5	09VXHS	09VXHS drosophila
62	36	85.7	371	10	09SXG6	09SXG6 oryza sativ
63	36	85.7	378	10	09SNJ2	09SNJ2 oryza sativ
64	36	85.7	542	5	09GUY4	09GUY4 penaeus jap
65	36	85.7	610	5	020798	020798 caenorhabd
66	36	85.7	629	13	093310	093310 xenopus lae
67	36	85.7	715	10	09EY76	09EY76 zea mays (m
68	36	85.7	795	3	09E524	09E524 neurospora
69	36	85.7	818	5	09V177	09V177 drosophila
70	36	85.7	945	3	09P313	09P313 cyphonectr
71	36	85.7	1119	4	09C0C3	09C0C3 homo sapien
72	36	85.7	1306	11	0923M3	0923M3 mus musculu
73	36	85.7	1484	4	09BXF3	09BXF3 homo sapien
74	35	83.3	181	5	026335	026335 drosophila
75	35	83.3	183	5	09W0C1	09W0C1 drosophila
76	35	83.3	192	5	09W0C2	09W0C2 drosophila
77	35	83.3	238	11	091W26	091W26 rattus rat
78	35	83.3	247	5	001646	001646 drosophila
79	35	83.3	265	4	09BR01	09BR01 homo sapien
80	35	83.3	267	10	09M1W7	09M1W7 arabidopsis
81	35	83.3	289	5	09N3A8	09N3A8 caenorhabd
82	35	83.3	300	5	022679	022679 caenorhabd
83	35	83.3	304	11	09CY95	09CY95 mus musculu
84	35	83.3	312	10	09M9S0	09M9S0 arabidopsis
85	35	83.3	315	2	09F0T6	09F0T6 edta-degrad
86	35	83.3	317	12	064812	064812 autographa
87	35	83.3	321	10	09SHP8	09SHP8 arabidopsis
88	35	83.3	327	11	0921D1	0921D1 mus musculu
89	35	83.3	337	10	09LPE6	09LPE6 arabidopsis

90	35	83.3	354	12	09H8B2	09H8b2 homo sapien
91	35	83.3	354	12	099H32	099H32 heliooverp
92	35	83.3	369	6	09TTR3	09Ttr3 equus caball
93	35	83.3	402	10	09AY92	09ay92 oryza sativ
94	35	83.3	412	4	09NMF5	09nwf5 homo sapien
95	35	83.3	426	10	09M254	09m254 arabidopsis
96	35	83.3	486	5	019863	019863 caenorhabdi
97	35	83.3	494	5	018295	018295 caenorhabdi
98	35	83.3	518	5	0961K4	0961k4 drosophila
99	35	83.3	548	4	014818	014818 homo sapien
100	35	83.3	548	4	0969H7	0969h7 homo sapien
101	35	83.3	571	4	015913	015913 homo sapien
102	35	83.3	588	11	P97318	P97318 mus musculu
103	35	83.3	597	3	092258	092258 podospora a
104	35	83.3	600	3	09C215	09c215 neotropora
105	35	83.3	623	4	014819	014819 homo sapien
106	35	83.3	625	11	064225	064225 mus musculu
107	35	83.3	638	4	092744	092744 homo sapien
108	35	83.3	639	4	015637	015637 homo sapien
109	35	83.3	639	4	092745	092745 homo sapien
110	35	83.3	639	11	092124	092124 mus musculu
111	35	83.3	653	11	008817	008817 mus musculu
112	35	83.3	688	10	023774	023774 craterostifig
113	35	83.3	713	5	09VHK3	09vbk3 drosophila
114	35	83.3	726	4	09BQNR	09bqn8 homo sapien
115	35	83.3	805	12	092431	092431 bombyx mori
116	35	83.3	916	5	09VT80	09vt80 drosophila
117	35	83.3	945	10	022196	022196 arabidopsis
118	35	83.3	952	5	017569	017569 caenorhabdi
119	35	83.3	1271	4	09UPN6	09upn6 homo sapien
120	35	83.3	1485	4	09UL15	09ul15 homo sapien
121	35	83.3	1691	5	09GRP7	09grp7 leishmania
122	35	83.3	1858	3	097B3	097b3 emeticeila
123	35	83.3	1957	4	09UPP3	09upp3 homo sapien
124	34	81.0	75	4	096FNR	096fn8 homo sapien
125	34	81.0	91	16	033059	033059 mycobacteri
126	34	81.0	107	17	059173	059173 pyrococcus
127	34	81.0	113	12	089348	089348 paramecium
128	34	81.0	117	4	09BRNR	09brnr8 homo sapien
129	34	81.0	119	12	069107	069107 herpes slimp
130	34	81.0	146	4	09NMAD	09nmd0 homo sapien
131	34	81.0	164	4	09FT76	09ft76 arabidopsis
132	34	81.0	185	6	09N0T7	09n0t7 bos taurus
133	34	81.0	195	11	099JNO	099j10 mus musculu
134	34	81.0	198	10	093Y15	093y15 capsella ru
135	34	81.0	205	2	09S2M0	09s2m0 streptomyce
136	34	81.0	211	2	049744	049744 mycobacteri
137	34	81.0	212	10	094JK9	094jk9 oryza sativ
138	34	81.0	213	6	09N0C4	09n0c4 macaca fasc
139	34	81.0	220	2	033130	033130 mycobacteri
140	34	81.0	226	5	020956	020956 caenorhabdi
141	34	81.0	236	11	0921Y0	0921y0 rattus norv
142	34	81.0	252	2	09REB0	09reb0 erwilia chr
143	34	81.0	252	10	09XE06	09xeu6 oryza sativ
144	34	81.0	252	10	09XF28	09xf28 oryza sativ
145	34	81.0	258	10	094H44	094h44 oryza sativ
146	34	81.0	267	2	0910H0	0910h0 streptomyce
147	34	81.0	268	13	09PMW3	09pmw3 brachydantio
148	34	81.0	275	10	09LW10	09lw10 arabidopsis
149	34	81.0	289	3	09HFL3	09hfl3 neosporea
150	34	81.0	293	10	09S809	09s809 oryza sativ

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	316 AA.
Q9VAX5				
AC	Q9VAX5:			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)			

DE	CG12872	PROTEIN.
GN	CG12872.	
OS	Drosophila melanogaster (Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Ephydroidea; Drosophilidae; Drosophila.	
OX	NCBI_TaxID=7227;	
RM	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BERKELEY.	
RX	MEDLINE=20196006; PubMed-10731132;	
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,	
RA	Amannatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,	
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	
RA	Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Chape M., Pfeiffer B.D.,	
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	
RA	Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,	
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	
RA	Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,	
RA	Botkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,	
RA	Buttis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,	
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,	
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	
RA	Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,	
RA	Postler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,	
RA	Glodex A., Gong F., Gorrell J.H., Gu Z., Guan F., Harris M.,	
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J., Houck J.,	
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,	
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,	
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,	
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,	
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,	
RA	Merulov G., Milshina N.V., Modary C., Morris J., Moshireli A.,	
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,	
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,	
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,	
RA	Rehner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,	
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,	
RA	Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,	
RA	Switskas R., Tector C., Turner R., Venter E., Wang A.B., Wang X.,	
RA	Wang Z.-Y., Wasserman D.A., Weinstein D.A., Weissbach J.,	
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,	
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,	
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,	
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,	
RT	"The genome sequence of Drosophila melanogaster."	
RL	Science 287:2185-2195(2000).	
DR	EMBL: AE003763; AAF56772.1; -	
DR	Flybase: FBgn0039569; CG12872.	
SQ	SEQUENCE 316 AA; 33984 MW; 2D50DD028DFE6901 CRC64;	

Query Match	100.0%;	Score 42;	DB 5;	Length 316;
Best local similarity	100.0%;	Pred. No. 11;		
Matches	7;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1 MOPPLP 7			
DB	249 MOPPLP 255			

RESULT	2	PRELIMINARY:	PRT:	1776 AA.
Q9MA20				
AC	Q9MA20:			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	T5E21.13.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shun P., Tanbunga G., Altafi H., Bel Q., Chin C., Chou J., Choi E.,
 RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mkharsky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Theoret A., Tortum J., Vaysberg M., Yu G., Federspiel N.A.,
 RA Theoret A., Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC T5E21 from chromosome
 1";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC010657; AAF63169.1; -
 DR HSSP: P02248; I0B1.
 DR InterPro: IPR002965; P_rich_extensn.
 DR InterPro: IPR000061; Surp.
 DR InterPro: IPR000626; ubiquitin.
 DR Pfam: PF01805; Surp; 4.
 DR Pfam: PF00240; ubiquitin; 2.
 DR PRINTS: PR01317; PRICHETENS.
 DR SMART: SM00213; UBQ; 1.
 DR PROSITE: PS50053; DBIQUITTIN_2; 1.
 SQ SEQUENCE 1776 AA; 200270 MW; C1C28169934E0A27 CRC64;

Query Match 95.2%; Score 40; DB 10; Length 1776;
 Best Local Similarity 85.7%; Pred. No. 1.le+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 :|||||:
 DB 1643 MOPPLP 1649

RESULT 3
 ID 09P38 PRELIMINARY; PRT; 612 AA.
 AC 09P38.
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE BETA-GALACTOSIDASE (EC 3.2.1.23) (LACTASE).
 GN XP0840.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OC NCBI_TaxID=2371;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-9A5C;
 RX MEDLINE-20365717; PubMed-10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvares R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferto J.A.,
 RA Fraga J.S., France S.C., Franco M.C., Frohne M., Furlan L.R.,
 RA Gantier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Kistner J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madela A.M.B.N., Madela H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhami A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,

RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siguelra W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Zeldanis J., Zetubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";
 RL Nature 406:151-159(2000).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING BETA-D-
 CC GALACTOSE RESIDUES IN BETA-D-GALACTOSIDES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 35 OF GLYCOSYL HYDROLASES.
 DR EMBL: AE003923; AAF83650.1; -
 DR InterPro: IPR001944; Glyco_hydro_35.
 DR Pfam: PF01301; Glyco_hydro_35; 1.
 DR PRINTS: PR00742; GLHYDRASE35.
 DR PROSITE: PS01182; GLYCOSYL_HYDROL_F35; 1.
 KW Complete proteome: Glycosylase, Hydrolase.
 SQ SEQUENCE 612 AA; 67893 MW; B894D6A3EB789BCB CRC64;

Query Match 92.9%; Score 39; DB 16; Length 612;
 Best Local Similarity 85.7%; Pred. No. 59;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 :|||||:
 DB 356 LOPPLP 362

RESULT 4
 ID 09GRR6 PRELIMINARY; PRT; 1404 AA.
 AC 09GRR6;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 152.9 KDA PROTEIN (FRAGMENT).
 GN L5882.01.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OC NCBI_TaxID=5664;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RA Masny D., Purnelle B., Goffeau A., Ivens A.C., Quail M.,
 RA Rajandream M.A., Barrell B.G., Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RX MEDLINE-98146435; PubMed-9477341;
 RA Ivens A.C., Lewis S.M., Bagherzaden A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the Leishmania major Friedlin genome";
 RL EMBL: AL45943; CAC14319.1; -
 DR EMBL: AL45943; CAC14319.1; -
 KW Hypothetical protein.
 FT NON-TER 1
 SQ SEQUENCE 1404 AA; 152867 MW; 25F42AB6702B9D3C CRC64;

Query Match 92.9%; Score 39; DB 5; Length 1404;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 :|||||:
 DB 1142 LOPPLP 1148

RESULT 5
 O9JLPO

```

ID 09JLP0 PRELIMINARY; PRT; 1444 AA.
AC 09JLP0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE REPEAT PROTEIN TYROSINE PHOSPHATASE-RHO.
GN PRPT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Besco J.A., Frostholt A., Popesco M.C., Burghes A.H.M., Rotter A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
DR EMBL; AF152556; AAD34158.4; -.
DR HSSP; P28827; IRPM.
DR MGP; MGI:1321152; PpPrt.
DR InterPro; IPR003962; PpIII_repeat.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003599; 19.
DR InterPro; IPR000998; MAM.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; TYR_prot_phptase.
DR Pfam; PF00041; fn3; 3.
DR Pfam; PF00629; MAM; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00020; MAMDOMAIN.
DR PRINTS; PR00700; PTPPHPTASE.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00409; IG; 1.
DR SMART; SM00137; MAM; 1.
DR SMART; SM00194; PTPc; 2.
DR PROSITE; PS00740; MAM_1; UNKNOWN_1.
DR PROSITE; PS00060; MAM_2; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN_2.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 2.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 2.
KM Glycoprotein; Repeat..
SQ SEQUENCE 1444 AA; 162218 MW; C7ECABA71FA21B82 CRC64;

Query Match
Best Local Similarity 92.9%; Score 39; DB 11; Length 1444;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPLP 7
DB 15 LQPPPLP 21

RESULT 6
Q9SGY5 PRELIMINARY; PRT; 542 AA.
ID Q9SGY5;
AC Q9SGY5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE P20B24.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shlin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bel O., Chin C., Chiu J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharisky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,

```

```

RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F20B24 from chromosome
RT I."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES.
CC (POLYGALACTURONASES).
DR EMBL; AC009398; AAF1670.1; -.
DR InterPro; IPR000743; Polygalacturonase.
DR Pfam; PF00295; Glyco_hydro_28; 2.
DR PROSITE; PS00502; POLYGALACTURONASE; 1.
KM Cell wall; Glycosidase; Hydrolase.
SQ SEQUENCE 542 AA; 59705 MW; 7C67D0C3BCD19BC8 CRC64;

Query Match
Best Local Similarity 90.5%; Score 38; DB 10; Length 542;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQPPLP 7
DB 105 VQPPPLP 111

RESULT 7
Q87610 PRELIMINARY; PRT; 72 AA.
ID Q87610;
AC Q87610;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE REV PROTEIN (FRAGMENT).
GN REV.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=111723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F031;
RX MEDLINE=97138325; PubMed=8985351;
RA Bibollet-Ruche F., Brengues C., Galat-Luong A., Galat G., Pourrut X.,
RA Vidal N., Veas F., Durand J.P., Cuny G.;
RT "Genetic diversity of simian immunodeficiency viruses from West
RT African green monkeys: evidence of multiple genotypes within
RT populations from the same geographical locale."
RL J. Virol. 71:307-313(1997).
DR EMBL; U37213; AAC56168.1; -.
DR InterPro; IPR000625; REV.
DR FT NON_TER 1 1
FT NON_TER 72 72
FT NON_TER 1 1
SQ SEQUENCE 72 AA; 8188 MW; B7A8027FA9BA432A CRC64;

Query Match
Best Local Similarity 88.1%; Score 37; DB 15; Length 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
DB 58 QPPPLP 63

RESULT 8
Q87603 PRELIMINARY; PRT; 73 AA.
ID Q87603;
AC Q87603;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE REV PROTEIN (FRAGMENT).
GN REV.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).

```

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-P058;
 RX MEDLINE=97138325; PubMed=8985351;
 RA Billolet-Ruche F., Brenques C., Galat-Luong A., Galat G., Pourrut X.,
 RA Vidal N., Veas F., Durand J.P., Cuny G.;
 RT "Genetic diversity of simian immunodeficiency viruses from West
 RT African green monkeys: evidence of multiple genotypes within
 RT populations from the same geographical locale.";
 RL J. Virol. 71:307-313(1997).
 DR EMBL: U37210; AAC56159.1;
 DR InterPro: IPR000625; REV.
 DR Pfam: PF00424; REV; 1.
 DR NON_TER 1
 FT SEQUENCE 73 AA; 8217 MW; A31FDD937DC9BB4D CRC64;
 SQ
 Query Match 88.1%; Score 37; DB 15; Length 73;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 OPPPLP 7
 Db 58 OPPPLP 63
 RESULT 9
 ID 087482 PRELIMINARY; PRT; 99 AA.
 AC 087482;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE REV.
 GN REV.
 OS Chimpanzee immunodeficiency virus (STV(cpz)) (CIV).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-D30;
 RX MEDLINE=96013851; PubMed=7474168;
 RA Jubler-Maurin V., Sarrail-Manchado P., Veas F., Vidal N.,
 RA Billolet-Ruche F., Durand J.P., Galat-Luong A., Cuny G.;
 RT "Regulatory genes of simian immunodeficiency viruses from west African
 RT green monkeys (Cercopithecus aethiops sabaeus).";
 RL J. Virol. 69:7349-7353(1995).
 DR EMBL: U20966; AAA91117.1;
 DR InterPro: IPR000625; REV.
 DR Pfam: PF00424; REV; 1.
 DR NON_TER 1
 FT SEQUENCE 99 AA; 11464 MW; 38AA788A294845E6 CRC64;
 SQ
 Query Match 88.1%; Score 37; DB 15; Length 99;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 OPPPLP 7
 Db 80 OPPPLP 85
 RESULT 10
 ID 09NDS3 PRELIMINARY; PRT; 126 AA.
 AC 09NDS3;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE DD9A (FRAGMENT).
 GN DD9A.

OS Pennaeus japonicus (Kuruma prawn).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
 OC Penaeidae; Marsupenaeus.
 OX NCBI_TaxID=27405;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EPIDERMIS;
 RX MEDLINE=20299686; PubMed=10840648;
 RA Watanabe T., Persson P., Endo H., Kono M.;
 RT "Molecular analysis of two genes, DD9A and B, which are expressed
 RT during the postmolt stage in the decapod crustacean Pennaeus
 RT japonicus.";
 RL Comp. Biochem. Physiol. 125B:127-136(2000).
 DR EMBL: AB031223; BA90875.1;
 DR InterPro: IPR000618; Insect_cuticle.
 DR Pfam: PF00379; Insect_cuticle; 1.
 DR PRINTS: PR00947; CUTICLE.
 DR PROSITE: PS00233; CUTICLE; UNKNOWN_1.
 DR NON_TER 1
 FT SEQUENCE 126 AA; 13809 MW; 0B39FECBB284C3B CRC64;
 SQ
 Query Match 88.1%; Score 37; DB 5; Length 126;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 OPPPLP 7
 Db 94 OPPPLP 99
 RESULT 11
 ID 09NDS2 PRELIMINARY; PRT; 126 AA.
 AC 09NDS2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE DD9B (FRAGMENT).
 GN DD9B.
 OS Pennaeus japonicus (Kuruma prawn).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
 OC Penaeidae; Marsupenaeus.
 OX NCBI_TaxID=27405;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-TAIL FAN;
 RX MEDLINE=20299686; PubMed=10840648;
 RA Watanabe T., Persson P., Endo H., Kono M.;
 RT "Molecular analysis of two genes, DD9A and B, which are expressed
 RT during the postmolt stage in the decapod crustacean Pennaeus
 RT japonicus.";
 RL Comp. Biochem. Physiol. 125B:127-136(2000).
 DR EMBL: AB031224; BA90876.1;
 DR InterPro: IPR000618; Insect_cuticle.
 DR Pfam: PF00379; Insect_cuticle; 1.
 DR PRINTS: PR00947; CUTICLE.
 DR NON_TER 1
 FT SEQUENCE 126 AA; 13787 MW; AA75D86F0042A67C CRC64;
 SQ
 Query Match 88.1%; Score 37; DB 5; Length 126;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 OPPPLP 7
 Db 94 OPPPLP 99
 RESULT 12

018043 PRELIMINARY; PRT; 164 AA.
 ID 018043;
 AC 018043;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE HYPOTHETICAL 17.1 KDA PROTEIN.
 GN C16B8.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peloderiinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Bentley D.;
 RT "The sequence of C. elegans cosmid C16B8.";
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U41031; AAA82620.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 164 AA; 17074 MW; 00EA97D08209A718 CRC64;

Query Match 88.1%; Score 37; DB 5; Length 164;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 Db 97 OPPPLP 102

RESULT 13
 ID 048681 PRELIMINARY; PRT; 178 AA.
 AC 048681;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, last sequence update)
 DT 01-AUG-1998 (TREMblrel. 07, last annotation update)
 DE F316.7 PROTEIN.
 GN F316.7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Federspiel N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,
 RA Au M., Arango R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,
 RA Oji O., Osborne B.I., Shinn P., Sun H., Toriumi M., Vysotskaia V.S.,
 RA Yu G., Ecker J., Theologis A., Davis R.W.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC002396; AAC00574.1; -.
 SQ SEQUENCE 178 AA; 19730 MW; 7A526DCB7F8423DD CRC64;

Query Match 88.1%; Score 37; DB 10; Length 178;

Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 Db 162 OPPPLP 167

RESULT 14
 ID 020796 PRELIMINARY; PRT; 186 AA.
 AC 020796;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE F55A11.1 PROTEIN.
 GN F55A11.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peloderiinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kershaw J.K.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z72511; CAA96655.1; -.
 DR InterPro: IPR002048; EF-hand.
 DR PROSITE: P500018; EF HAND; UNKNOWN 2.
 SQ SEQUENCE 186 AA; 20745 MW; 01A61CB630216D1 CRC64;

Query Match 88.1%; Score 37; DB 5; Length 186;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 Db 127 OPPPLP 132

RESULT 15
 ID 091653 PRELIMINARY; PRT; 294 AA.
 AC 091653;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE HYPOTHETICAL PROTEIN.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RA "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:p0699011.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP002817; BAB03438.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 294 AA; 31954 MW; 94663EDBF8183B35 CRC64;

Query Match 88.1%; Score 37; DB 10; Length 294;

Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OPPPLP 7
DB 144 OPPPLP 149

RESULT 16

Q9BSM2 PRELIMINARY; PRT; 322 AA.
AC Q9BSM2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE HYPOTHETICAL 34.9 KDA PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIMPH, LYMPHOMA;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004950; AA04950.1;
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 322 AA; 34916 MW; 33C65E86B8F9D761 CRC64;

Query Match 88.1%; Score 37; DB 4; Length 322;

Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OPPPLP 7
DB 6 OPPPLP 11

RESULT 17

Q99LN2 PRELIMINARY; PRT; 346 AA.
AC Q99LN2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE HYPOTHETICAL 39.3 KDA PROTEIN (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002302; AA02302.1;
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 346 AA; 39305 MW; 2CB318DECTEDAE21 CRC64;

Query Match 88.1%; Score 37; DB 11; Length 346;

Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OPPPLP 7
DB 299 OPPPLP 304

RESULT 18

Q96RZ4

ID Q96RZ4 PRELIMINARY; PRT; 357 AA.

AC Q96RZ4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE HYPOTHETICAL 38.3 KDA PROTEIN.
GN GS93.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21096910; PubMed=11157797;
RA Dantels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA Higgs D.R.;
RT *Sequence, structure and pathology of the fully annotated terminal 2
RT MD of the short arm of human chromosome 16.*;
RL Hum. Mol. Genet. 10:339-352(2001).
DR EMBL; AE006466; AA61273.1;
KW Hypothetical protein.
SQ SEQUENCE 357 AA; 38342 MW; A8EC29C06E5C357B CRC64;

Query Match 88.1%; Score 37; DB 4; Length 357;

Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OPPPLP 7
DB 30 OPPPLP 35

RESULT 19

Q22921 PRELIMINARY; PRT; 393 AA.
AC Q22921;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE PUTATIVE WRKY-TYPE DNA-BINDING PROTEIN (WRKY TRANSCRIPTION FACTOR
DE 25).
GN AT2G30250 OR WRKY25.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayan L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT *Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.*;
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=INFLORESCENSE;
RA Ulker B., Kuschlir S., Somsich I.E.;
RT *Arabidopsis thaliana transcription factor WRKY25.*;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC002338; AAC16930.1; -
 DR EMBL: AF418309; AAL13040.1; -
 DR InterPro: IPR003657; WRKY.
 DR Pfam: PF03106; WRKY; 2.
 KW DNA-binding.
 SQ SEQUENCE 393 AA; 44133 MW; 9057D65B69E12A0C CRC64;

Query Match 88.1%; Score 37; DB 10; Length 393;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 |||||
 DB 50 QPPPLP 55

RESULT 20
 09PTS7 PRELIMINARY; PRT; 421 AA.
 AC 09PTS7;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE P0409B08.10 PROTEIN (P0044F08.28 PROTEIN).
 GN P0409B08.10 OR P0044F08.28.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza.
 NX NCBI_TaxID=4530;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0409B08.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0044F08.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP002860; BAB18283.1; -
 DR EMBL: AP002909; BAB21193.1; -
 DR InterPro: IPR002965; P-rich_extensn.
 DR PRINTS: PR01217; PRICHEXTENS.
 SQ SEQUENCE 421 AA; 44905 MW; 8D3BD1A78789F55D CRC64;

Query Match 88.1%; Score 37; DB 10; Length 421;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 |||||
 DB 116 QPPPLP 121

RESULT 21
 09V6M1 PRELIMINARY; PRT; 440 AA.
 AC 09V6M1;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CG4663 PROTEIN.
 GN CG4663.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NX NCBI_TaxID=7227;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Chape M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Ewanellista C.C., Ferraz C., Ferlita S., Fleischmann W.,
 RA Fostler A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshireli A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Welstock G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AEO03819; AAP58402.1; -
 DR Flybase: FBgn0033812; CG4663.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00018; SH3; 1.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50002; SH3; 1.
 SQ SEQUENCE 440 AA; 46660 MW; 3ABEC719DDE9C50D CRC64;

Query Match 88.1%; Score 37; DB 5; Length 440;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 |||||
 DB 55 QPPPLP 60

RESULT 22
 09N906 PRELIMINARY; PRT; 456 AA.
 AC 09N906;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE POSSIBLE SURFACE ANTIGEN PROTEIN.
 GN L1177.05.
 OS Leishmania major.

OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA Hilbert H., Wedler H., Wedler E., Duesterhoeft A., Ivens A.C.,
 RA Quall M., Rajandream M.A., Barrell B.G.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RL "A physical map of the Leishmania major Friedlin genome.";
 RL Genome Res. 8:135-145(1998).
 DR EMBL: AL354532; CAB89607.1; -.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003592; LRR_out.
 DR Pfam: PF00560; LRR; 9.
 DR SMART: SM00370; LRR; 7.
 SQ SEQUENCE 456 AA; 50097 MW; DA17AB58634478CD CRC64;

Query Match 88.1%; Score 37; DB 5; Length 456;
 Best Local Similarity 85.7%; Pred. No. 97;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 DB 1 MOPPLP 7

RESULT 23
 OY 093571 PRELIMINARY; PRT; 471 AA.
 AC 093571;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE F25H2.8 PROTEIN.
 GN F25H2.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilkinson J.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL: Z79754; CAB82096.1; -.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR000608; UBQ_conjugat.
 DR SMART: SM00212; UBCc; 1.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 DR PROSITE: PS00127; UBOUTIN_CONJUGAT; 2; 1.
 SQ SEQUENCE 471 AA; 54029 MW; 69EBBCA5B510D78 CRC64;

Query Match 88.1%; Score 37; DB 5; Length 471;
 Best Local Similarity 71.4%; Pred. No. 1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOPPLP 7
 DB 451 LQPPPIP 457

RESULT 24
 OY 022717 PRELIMINARY; PRT; 588 AA.
 AC 022717;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE SERINE/THREONINE PROTEIN KINASE.
 GN F8A5.31.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Federspiel N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,
 RA Au M., Araujo R., Brendel V., Buehler E., Dewar K., Feng J., Kim C.,
 RA Li Y., Qi O., Osborne B.I., Shinn P., Sun H., Toriumi M.,
 RA Yatsushita V.S., Yu G., Ecker J., Theologis A., Davis R.W.;
 RT "Genomic sequence of Arabidopsis.";
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AC002292; AAB71968.1; -.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR004040; STY_pkinase.
 DR Pfam: PF00560; LRR; 3.
 DR Pfam: PF00069; pkinase; 1.
 DR SMART: SM00370; LRR; 2.
 DR SMART: SM00221; STYc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 588 AA; 65462 MW; CCE9D5DBA4B03D138 CRC64;

Query Match 88.1%; Score 37; DB 10; Length 588;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 DB 550 QPPPLP 555

RESULT 25
 OY 099NA4 PRELIMINARY; PRT; 629 AA.
 AC 099NA4;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE IKAPPAB-ZETA.
 GN MAIL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21347944; PubMed=11356851;
 RA Yamazaki S., Muta T., Takeshige K.;
 RT "A Novel Ikappab Protein, Ikappab-zeta, Induced by Proinflammatory
 RT Stimuli, Negatively Regulates Nuclear Factor-kappab in the Nuclei.";
 RL J. Biol. Chem. 276:27657-27662(2001).
 DR EMBL: AB047549; BAB32782.1; -.
 DR HSP; P25963; INFI.

DR MGD; MG1:1931595; Mail.
 DR InterPro: IPR002110; ANK.
 DR PRINTS: PR01415; ANKRYIN.
 DR SMART: SM00248; ANK; 6.
 DR PROSITE: PSS0088; ANK_REPEAT; 3.
 DR PROSITE: PSS0297; ANK_REPEAT_REGION; 1.
 DR ANK repeat; Repeat.
 KW ANK repeat; Repeat.
 SO SEQUENCE 629 AA; 69179 MW; 8F87A6147324534 CRC64;

Query Match 88.1%; Score 37; DB 11; Length 629;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 |||||
 Db 197 QPPPLP 202

RESULT 26

O93ZS4

AC 093ZS4; PRELIMINARY; PRT; 632 AA.

DT 01-DEC-2001 (TREMblrel. 19, Created)

DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DE PUTATIVE SERINE/THREONINE PROTEIN KINASE.

GN F8A5.31.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,

RA Onodera C.S., Quach H.L., Tang C., Tortum M., Yamamura Y., Yu G.,

RA Yu S., Bowser L., Carrincci P., Chen H., Cheuk R., Hayashizaki Y.,

RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,

RA Koeseema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,

RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,

RA Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,

RA Theologis A.;

RT "Full length cDNA of gene F8A5.31 (GI:2462749)."

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY056294; AAL07143.1;

KW Kinase.

SO SEQUENCE 632 AA; 70171 MW; 7DC54B2707188B8C CRC64;

Query Match 88.1%; Score 37; DB 10; Length 632;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 |||||
 Db 594 QPPPLP 599

RESULT 27

O9DFZ5

AC 09DFZ5; PRELIMINARY; PRT; 651 AA.

DT 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DE HYPOHETICAL 71.6 KDA PROTEIN.

GN DKFZP34D064.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-TESTIS;

RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL10225; CAB53683.1;

DR InterPro: IPR002108; Cofilin_ADF.

DR Pfam: PF00241; Cofilin_ADF; 1.

DR SMART: SM00102; ADF; 1.

KW Hypothetical protein.

SO SEQUENCE 651 AA; 71594 MW; 8F7600A6A9E96C1 CRC64;

Query Match 88.1%; Score 37; DB 4; Length 651;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 |||||
 Db 365 QPPPLP 370

RESULT 28

O35745

AC 035745; PRELIMINARY; PRT; 667 AA.

DT 01-JAN-1998 (TREMblrel. 05, Created)

DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)

DE ZINC FINGER PROTEIN REGULATOR OF APOPTOSIS AND CELL CYCLE ARREST.

GN ZAC1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BALB/C; TISSUE-PITUITARY GLAND;

RX MEDLINE-97327559; PubMed-9184226;

RA Spengler D., Villalba M., Hoffmann A., Pantaloni C., Houssami S.,

RA Bockert J., Journot L.;

RT "Regulation of apoptosis and cell cycle arrest by Zac1, a novel zinc

finger protein expressed in the pituitary gland and the brain."

RL EMBL J. 16:2814-2825(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BALB/C; TISSUE-PITUITARY GLAND;

RA Spengler D., Villalba M., Hoffmann A., Pantaloni C., Houssami S.,

RA Bockert J., Journot L.;

RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

DR EMBL; X95503; CAA64757.1;

DR EMBL; X95504; CAA64758.1;

DR MGD; MG1:1100874; Zac1.

DR InterPro: IPR000822; Znf-C2H2.

DR Pfam: PF00096; Zf-C2H2; 7.

DR PRINTS: PR00048; ZINC_FINGER.

DR SMART: SM00355; ZNF_C2H2; 7.

DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 7.

DR PROSITE: PSS0157; ZINC_FINGER_C2H2_2; 6.

KW DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.

SO SEQUENCE 667 AA; 75258 MW; 392F0AC979BC10F CRC64;

Query Match 88.1%; Score 37; DB 11; Length 667;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 |||||
 Db 527 QPPPLP 532

RESULT 29

```

09EPT3
ID 09EPT3 PRELIMINARY; PRT; 675 AA.
AC 09EPT3:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE LOST ON TRANSFORMATION PROTEIN 1.
GN ZAC1 OR LOT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathl; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NIH-SWISS;
RX MEDLINE=20065118; PubMed=10597250;
RA Abdollahi A., Bao R., Hamilton T.C.;
RT *LOT1 is a growth suppressor gene down-regulated by the epidermal
RT growth factor receptor ligands and encodes a nuclear zinc-finger
RT protein.*;
RL Oncogene 18:6477-6487(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NIH-SWISS;
RA Abdollahi A., Hamilton T.C.;
RT *Identification of a splice variant of mouse lot1 gene.*;
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF324471; AAF48331.1; -.
DR MGD; MGI:1100874; Zacl.
DR InterPro; IPR000822; Znf-C2H2.
DR Pfam; PF00096; znf-C2H2; 7.
DR SMART; SM00355; znf-C2H2; 7.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 6.
DR DNA-binding; Metal-binding; Zinc-finger.
KW SEQUENCE 675 AA; 76109 MW; 0F467D165CA5B097 CRC64;
SQ

Query Match 88.1%; Score 37; DB 11; Length 675;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
DB 524 QPPPLP 529

RESULT 30
044760 PRELIMINARY; PRT; 692 AA.
AC 044760:
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-OCT-2001 (TREMBLrel. 18, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE HYPOHETICAL 76.8 KDA PROTEIN.
GN T12F5.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT *Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.*;
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Clarke K., Wohldmann P., Rohlfing T., Bauer C.;
RT *The sequence of C. elegans cosmid T12F5.*;

```

```

RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT *Direct Submission.*;
RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF039718; AAB96747.2; -.
DR Hypothetical protein.
KW SEQUENCE 692 AA; 76832 MW; 26C8DB21B8A35782 CRC64;
SQ

Query Match 88.1%; Score 37; DB 5; Length 692;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
DB 45 QPPPLP 50

RESULT 31
09JLQ4 PRELIMINARY; PRT; 704 AA.
AC 09JLQ4:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE ZINC FINGER PROTEIN ZAC1.
GN ZAC1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathl; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20136063; PubMed=10669760;
RA Huang S.-M., Stallcup M.R.;
RT *Mouse Zacl, a Transcriptional Coactivator and Repressor for Nuclear
RT Receptors.*;
RL Mol. Cell. Biol. 20:1855-1867(2000).
DR EMBL; AF147785; AAF34245.1; -.
DR MGD; MGI:1100874; Zacl.
DR InterPro; IPR000822; Znf-C2H2.
DR Pfam; PF00096; znf-C2H2; 7.
DR SMART; SM00355; znf-C2H2; 7.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 6.
DR DNA-binding; Metal-binding; Zinc-finger.
KW SEQUENCE 704 AA; 79214 MW; 3AF88262D504FA80 CRC64;
SQ

Query Match 88.1%; Score 37; DB 11; Length 704;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
DB 527 QPPPLP 532

RESULT 32
09AC37 PRELIMINARY; PRT; 721 AA.
AC 09AC37:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE AFI605230/YUP8H12.16.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

```

OX NCBI_TaxID-3702;
 [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Shinn P.,
 RA Treacy S.E., Banh J., Bowser L., Carinaci P., Chung M.K.,
 RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamuya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L.,
 RA Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
 RA Toriumi M., Yamada K., Yu G., Shinzaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones";
 RL Submitted (Mar-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY031177; AAK59762.1;
 SO SEQUENCE 721 AA; 79423 MW; 652287B59AFCB9 CRC64;

Query Match 88.1%; Score 37; DB 10; Length 721;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 QPPPLP 7
 I I I I I
 DB 205 MSPPPLP 211

RESULT 33
 Q9EST8 PRELIMINARY; PRT; 728 AA.
 AC Q9EST8;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MAIL (R1-1 INDUCIBLE NUCLEAR ANKYRIN-REPEAT PROTEIN).
 GN MAIL OR MAIL OR INAP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID-10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-SPLEEN;
 RX MEDLINE-20539924; PubMed-11086164;
 RA Kitamura H., Kanenita K., Okita K., Morimatsu M., Saito M.;
 RT "Mail, a novel nuclear IB protein that potentiates LPS-induced IL-6
 production.";
 RL FEBS Lett. 485:53-56(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21201066; PubMed-11278262;
 RA Haruta H., Kato A., Todokoro K.;
 RT "Isolation of a Novel Interleukin-1-Inducible Nuclear Protein Bearing
 Ankyrin-repeat Motifs.";
 RL J. Biol. Chem. 276:12485-12488(2001).
 DR EMBL: AB020974; BAB18302.1;
 DR EMBL: AB026551; BAA95161.2;
 DR HSSP; P25963; INFI.
 DR MGD; MGI:1931595; Mail.
 DR InterPro; IPR002110; ANK.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 6.
 DR PROSITE; PS00086; ANK_REPEAT; 3.
 DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
 DR ANK repeat; Repeat.
 SO SEQUENCE 728 AA; 79006 MW; E2CAB725295A08B7 CRC64;

Query Match 88.1%; Score 37; DB 11; Length 728;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 I I I I I

DB 296 QPPPLP 301

RESULT 34
 ID 096S74 PRELIMINARY; PRT; 731 AA.
 AC 096S74;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE RHO-GTPASE ACTIVATING PROTEIN.
 GN ARHGAP9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID-9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21290655; PubMed-11396949;
 RA Furukawa Y., Kawsoc T., Daigo Y., Nishiwaki T., Ishiguro H.,
 RA Takahashi M., Kitayama J., Nakamura Y.;
 RT "Isolation of a novel human gene, ARHGAP9, encoding a rho-GTPase
 activating protein.";
 RT Biochem. Biophys. Res. Commun. 284:643-649(2001).
 DR EMBL: AB051853; BAB56159.1;
 SO SEQUENCE 731 AA; 81150 MW; B63901EACA109D02 CRC64;

Query Match 88.1%; Score 37; DB 4; Length 731;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 I I I I I
 DB 131 QPPPLP 136

RESULT 35
 Q922V1 PRELIMINARY; PRT; 733 AA.
 ID Q922V1;
 AC Q922V1;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE UNKNOWN (PROTEIN FOR IMAGE:3709003) (FRAGMENT).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID-10090;
 [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC006755; AAH06755.1;
 FT NON TER 1
 SO SEQUENCE 733 AA; 84405 MW; 6772130812C7A10A CRC64;

Query Match 88.1%; Score 37; DB 11; Length 733;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 I I I I I
 DB 704 QPPPLP 709

RESULT 36
 ID 091Z55 PRELIMINARY; PRT; 738 AA.
 AC 091Z55;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE HYPOThETICAL 85.0 KDA PROTEIN (FRAGMENT).

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RN [1]
RP SEQUENCE FROM N.A.

RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC009668; AA09668.1; -
KW Hypothetical protein.

FT NON_TER
SQ SEQUENCE 738 AA; 84974 MW; E0FC99A1FF233P56 CRC64;

Query Match 88.1%; Score 37; DB 11; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
DB 709 QPPPLP 714

RESULT 37

ID 023045 PRELIMINARY; PRT; 749 AA.

AC 023045;
DT 01-JAN-1998 (TREMblrel. 05, Created)

DT 01-JUN-1998 (TREMblrel. 05, last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, last annotation update)

DE YUP8H12.16 PROTEIN.

GN YUP8H12.16.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]
RP SEQUENCE FROM N.A.

RA Strausberg R.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; Y00000; AA00000.1; -

DR InterPro: IPR001356; Homeobox.

DE InterPro: IPR002913; STARF.

DR Pfam: PF00046; homeobox; 2.

DR Pfam: PF01852; STARF; 1.

DR SMART; SM00344; HOX; 1.

DR PROSITE; PS00027; HOMEBOX.1; 1.

DR PROSITE; PS00071; HOMEBOX.2; 1.

KN DNA-binding; Homeobox; Nuclear protein.

SQ SEQUENCE 749 AA; 83038 MW; AE180DBEE394E25 CRC64;

Query Match

88.1%; Score 37; DB 10; Length 749;

Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MQPPPLP 7
DB 221 MSPPPLP 227

RESULT 38

ID 09BR9 PRELIMINARY; PRT; 750 AA.

AC 09BR9;
DT 01-JUN-2001 (TREMblrel. 17, Created)

DT 01-JUN-2001 (TREMblrel. 17, last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, last annotation update)

DE HYPOThETICAL 83.3 KDA PROTEIN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.

RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC006107; AA06107.1; -

DR HSP; P07751; ITUC.

DR InterPro: IPR001849; PH.

DR InterPro: IPR000198; RhogAP.

DR InterPro: IPR001452; SH3.

DR Pfam: PF00169; PH; 1.

DR Pfam: PF00620; RhogAP; 1.

DR Pfam: PF00018; SH3; 1.

DR SMART; SM00233; PH; 1.

DR SMART; SM00324; RhogAP; 1.

DR SMART; SM00326; SH3; 1.

DR SMART; SM00456; WW; 1.

DR PROSITE; PS50003; PH DOMAIN; 1.

DR PROSITE; PS50002; SH3; 1.

DR PROSITE; PS50020; WW DOMAIN; 2; 1.

DR Hypothetical protein.

SQ SEQUENCE 750 AA; 83293 MW; 0603BAEE4043B315 CRC64;

Query Match 88.1%; Score 37; DB 4; Length 750;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
DB 131 QPPPLP 136

RESULT 39

ID 09VHR3 PRELIMINARY; PRT; 819 AA.

AC 09VHR3;
DT 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, last sequence update)

DT 01-MAY-2000 (TREMblrel. 13, last annotation update)

DE CG11718 PROTEIN.

GN CG11718.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]
RP SEQUENCE FROM N.A.

RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; Y00000; AA00000.1; -

DR InterPro: IPR001356; Homeobox.

DE InterPro: IPR002913; STARF.

DR Pfam: PF00046; homeobox; 2.

DR Pfam: PF01852; STARF; 1.

DR SMART; SM00344; HOX; 1.

DR PROSITE; PS00027; HOMEBOX.1; 1.

DR PROSITE; PS00071; HOMEBOX.2; 1.

KN DNA-binding; Homeobox; Nuclear protein.

SQ SEQUENCE 749 AA; 83038 MW; AE180DBEE394E25 CRC64;

Query Match

88.1%; Score 37; DB 10; Length 749;

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abtill J.F., Agbayan A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Bailew R.M., Baas A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottler P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu N.E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugen-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hoslin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
 RA Jaitelli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 RA Palazolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003678; AAF54238.1; -;
 DR Flybase: FBgn0037585; CG11718.
 SQ SEQUENCE 819 AA; 92141 MW; E7F295E74FE2A72B CRC64;

Query Match 88.1%; Score 37; DB 5; Length 819;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 OPPPLP 7
 Db 125 OPPPLP 130
 RESULT 40
 093176 PRELIMINARY; PRT; 830 AA.
 AC 093176;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE C02C6.1A PROTEIN.
 GN C02C6.1A.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodermata; Caenorhabditis.
 OX NCBI_TaxID=6239;
 OX 1)
 RN SEQUENCE FROM N.A.
 RP Swinburne J.;
 RL Submitted (Aug-1996) to the EMBL/Genbank/DBJ databases.
 RN 12)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C. elegans*: A platform for

RT Investigating biology *;
 RL Science 282:2012-2018(1998).
 DR EMBL: Z79596; CAB01857.1; -;
 DR HSSP: Q05193; 2DYN.
 DR InterPro: IPR001401; Dynamlin.
 DR InterPro: IPR000375; Dynamlin_central.
 DR InterPro: IPR001330; GED.
 DR InterPro: IPR001849; PH.
 DR Pfam: PF00350; dynamlin_1.
 DR Pfam: PF01031; dynamlin_2.
 DR Pfam: PF02212; GED; 1.
 DR Pfam: PF00169; PH; 1.
 DR PRINTS: PR00195; DYNAMIN.
 DR SMART: SM00053; DYNCL; 1.
 DR SMART: SM00302; GED; 1.
 DR SMART: SM00233; PH; 1.
 DR PROSITE: PS00410; DYNAMIN; 1.
 DR PROSITE: PS0003; PH_DOMAIN; 1.
 SQ SEQUENCE 830 AA; 93407 MW; FF681250E51AB8A5 CRC64;

Query Match 88.1%; Score 37; DB 5; Length 830;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 OPPPLP 7
 Db 747 OPPPLP 752

RESULT 41
 090919 PRELIMINARY; PRT; 838 AA.
 AC 090919;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE DYNAMIN.
 GN DYN-1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodermata; Caenorhabditis.
 OX NCBI_TaxID=6239;
 OX 1)
 RN SEQUENCE FROM N.A.
 RP STRAIN-N2;
 RC MEDLINE=97439883; PubMed=9294229;
 RA Clark S.G., Shurland D.L., Meyerowitz E.M., Bargmann C.I.,
 RA Clark S.G., Shurland D.L., Meyerowitz E.M., Bargmann C.I.,
 RA van der Bliek A.M.;
 RT Inducible locomotion defect in *C. elegans*.;
 RT Proc. Natl. Acad. Sci. U.S.A. 94:10438-10443(1997).
 RL 12)
 RN SEQUENCE FROM N.A.
 RP STRAIN-N2;
 RA Clark S.G., Shurland D.L., Meyerowitz E.M., Bargmann C.I.,
 RA van der Bliek A.M.;
 RL Submitted (Jul-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF167982; AAD50438.1; -;
 DR HSSP: Q05193; 2DYN.
 DR InterPro: IPR001401; Dynamlin.
 DR InterPro: IPR000375; Dynamlin_central.
 DR InterPro: IPR001330; GED.
 DR InterPro: IPR001849; PH.
 DR Pfam: PF00350; dynamlin_1.
 DR Pfam: PF01031; dynamlin_2.
 DR Pfam: PF02212; GED; 1.
 DR Pfam: PF00169; PH; 1.
 DR PRINTS: PR00195; DYNAMIN.
 DR SMART: SM00053; DYNCL; 1.
 DR SMART: SM00302; GED; 1.
 DR SMART: SM00233; PH; 1.
 DR PROSITE: PS00410; DYNAMIN; 1.

DR PROSITE: PS50003; PH_DOMAIN; 1.
SQ SEQUENCE 838 AA; 94362 MW; 7A58B3CC2F7D16A3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 37; DB 5; Length 838;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
DB 747 OPPPLP 752

RESULT 42

ID 0950Y9 PRELIMINARY; PRT; 838 AA.
AC 0950Y9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE C02C6.1B PROTEIN.
CN C02C6.1B.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Swaburne J.;
RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
RM [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RT Science 282:2013-2018(1998).
RL EMBL; 279596; CAC42251.1;
SQ SEQUENCE 838 AA; 94421 MW; 78F89B6EA555B689 CRC64;

Query Match
Best Local Similarity 100.0%; Score 37; DB 5; Length 838;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
DB 747 OPPPLP 752

RESULT 43

ID 09NFD4 PRELIMINARY; PRT; 926 AA.
AC 09NFD4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 6209.1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Murphy L., Harris D., Barrell B.;
RT "Sequencing the distal X chromosome of Drosophila melanogaster."
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RM [2]
RP SEQUENCE FROM N.A.
RA Benos P.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
EMBL; AL009171; CAA15646.1;

DR InterPro: IPR000633; Vinculin_2.
DR InterPro: IPR001043; vinculin_acaten.
DR Pfam: PF01044; Vinculin; 1.

DR PRINTS: PR00806; VINCULIN.
DR PRODOM: PD025583; Vinculin_2; 2.
DR PROSITE: PS00664; VINCULIN_2; UNKNOWN; 1.
SQ SEQUENCE 926 AA; 102431 MW; 51A3E65A3C45987 CRC64;

Query Match
Best Local Similarity 100.0%; Score 37; DB 5; Length 926;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
DB 690 OPPPLP 695

RESULT 44

ID 046037 PRELIMINARY; PRT; 961 AA.
AC 046037;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE VINC PROTEIN.
GN VINC OR EG:103B4.1 OR CG3299.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Plankkuch C., Baldwin D.,
RA Bailley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart M.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeigbam C.,
RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mitsuhashi N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Murphy L., Harris D., Barrell B.;
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Benos P.;
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AE003423; AAF45752.1; -
 DR EMBL: AL009193; CA15691.1; -
 DR FlyBase: FBgn004397; Vinc.
 DR InterPro: IPR00633; Vinculin_2.
 DR Pfam: PF01044; Vinculin_1.
 DR PRINTS: PR00806; VINCULIN.
 DR PRODOM: PD025583; VINCULIN_2; 2; -
 DR PROSITE: PS00664; VINCULIN_2; 1; 17E50E11F507C875 CRC64;
 SQ SEQUENCE 961 AA; 106301 MW; 17E50E11F507C875 CRC64;

Query Match 88.1%; Score 37; DB 5; Length 961;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 Db 725 QPPPLP 730

RESULT 45
 ID 024584 PRELIMINARY; PRT; 962 AA.
 AC 024584;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE VINCULIN.
 GN VINC OR EG:10384.1 OR CG3299.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OC NCBI_Taxid=7227;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CT W(A).
 RA MEDLINE-97424359; PubMed-9280281;
 RX Alatorsev V., Kramerova I., Frolov M.V., Lavrov S.A., Westpahl E.D.;
 RT "Vinculin gene is non-essential in Drosophila melanogaster.";
 RL FEBS Lett. 413:197-201(1997).
 DR EMBL: X96601; CA65421.1; -
 DR FlyBase: FBgn004397; Vinc.
 DR InterPro: IPR00633; Vinculin_2.
 DR InterPro: IPR001043; Vinculin_acaten.
 DR Pfam: PF01044; Vinculin_1.
 DR PRINTS: PR00806; VINCULIN.
 DR PRODOM: PD025583; VINCULIN_2; 2.
 DR PROSITE: PS00664; VINCULIN_2; 1.
 SQ SEQUENCE 962 AA; 106510 MW; 1066FAA5856F1A9E CRC64;

Query Match 88.1%; Score 37; DB 5; Length 962;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QPPPLP 7
 Db 725 QPPPLP 730

RESULT 46
 OY3203

ID 093203 PRELIMINARY; PRT; 1125 AA.
 AC 093203;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE C11B4.6 PROTEIN.
 GN C11B4.6
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OC NCBI_Taxid=6239;
 OX [1]
 RN SEQUENCE FROM N.A.
 RA Mortimore B.J.;
 RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99069613; PubMed-9851916;
 RA none;
 RT Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.
 RL Science 282:2012-2018(1998).
 DR EMBL: Z81015; CAB02658.1; -
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000050; PID_domain.
 DR InterPro: IPR001660; SAM.
 DR Pfam: PF00023; ank; 6.
 DR Pfam: PF00536; SAM; 1.
 DR SMART: SMO0248; ANK; 4.
 DR SMART: SMO0454; SAM; 2.
 DR PROSITE: PS50088; ANK_REPEAT; 4.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE: PS01179; PID; 1.
 KW ANK repeat; Repeat.
 SQ SEQUENCE 1125 AA; 127133 MW; A9E7415A62FAD80F CRC64;

Query Match 88.1%; Score 37; DB 5; Length 1125;
 Best Local Similarity 85.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MQPPPLP 7
 Db 987 MQPPPLP 993

RESULT 47
 ID 013088 PRELIMINARY; PRT; 1154 AA.
 AC 013088;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ZEB (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_Taxid=9606;
 OX [1]
 RN SEQUENCE FROM N.A.
 RA Genetta T., Ruzitsky D., Kadesch T.;
 RT "Displacement of an E-box-binding repressor by basic-helix-loop-helix
 RT proteins: Implications for B-cell specificity of the immunoglobulin
 RT heavy-chain enhancer.";
 RL Mol. Cell. Biol. 15:0-0(1995).
 DR EMBL: U19969; AAA62155.1; -
 DR TRANSFAC: T00625; -
 DR TRANSFAC: T02324; -
 DR InterPro: IPR001356; Homeobox.
 DR InterPro: IPR000822; Znf-C2H2.
 DR Pfam: PF00096; zf-C2H2; 7.
 DR SMART: SMO0389; HOX; 1.
 DR SMART: SMO0355; Znf_C2H2; 7.

DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 5.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 6.
 KW DNA-binding; Metal-binding; zinc-finger.
 FT NON_TER 1
 SQ SEQUENCE 1154 AA; 126924 MW; 1AADAC7851909113 CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 |||||
 Db 591 OPPPLP 596

RESULT 48
 O9JKT1 PRELIMINARY; PRT; 1179 AA.
 AC O9JKT1;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SCARFOLDING PROTEIN SLIPR.
 GN SLIPR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Arnold E., Adamsky K., Peles E.;
 RT "SLIPR, a novel scaffolding molecule that interacts with receptor
 RT protein tyrosine phosphatase b (RPTPb).";
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF255614; AAF66069.1; -.
 DR HSSP: Q12923; 3PDZ.
 DR InterPro: IPR000619; Guanylate_kin.
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR001202; MW.
 DR Pfam: PF00625; Guanylate_kin; 1.
 DR Pfam: PF00595; PDZ; 5.
 DR SMART: SM0072; GUKC; 1.
 DR SMART: SM00228; PDZ; 6.
 DR SMART: SM00456; MW; 2.
 DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
 DR PROSITE: PS0052; GUANYLATE_KINASE_2; 1.
 DR PROSITE: PS50106; PDZ; 6.
 DR PROSITE: PS01159; MW_DOMAIN_1; 2.
 DR PROSITE: PS00020; MW_DOMAIN_2; 2.
 SQ SEQUENCE 1179 AA; 129404 MW; BCC29F3434E0C1F CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 |||||
 Db 1128 OPPPLP 1133

RESULT 49
 O9QY22 PRELIMINARY; PRT; 1263 AA.
 AC O9QY22;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 144.1 KDA PROTEIN (FRAGMENT).
 GN RAB61P1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ICR OUTBRED STRAIN;
 RX MEDLINE-21108749; PubMed-11166556;
 RA Callebaut I., De Gunzburg J., Goud B., Morrison J.P.;
 RT "RUN domains: a new family of domains involved in ras-like GTPase
 RT signaling";
 RL Trends Biochem. Sci. 26:79-83(2001).
 DR EMBL: AJ245569; CAB55599.1; -.
 DR MGD: MGI:1201681; Rab61P1.
 DR InterPro: IPR001194; DENN.
 DR InterPro: IPR001024; LH2.
 DR InterPro: IPR004012; Run.
 DR Pfam: PF02141; DENN; 1.
 DR Pfam: PF01477; PLAT; 1.
 DR Pfam: PF02759; RUN; 1.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 1263 AA; 144126 MW; 03818E475B6A3758 CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 |||||
 Db 210 OPPPLP 215

RESULT 50
 O9JFV0 PRELIMINARY; PRT; 1291 AA.
 AC O9JFV0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 147.4 KDA PROTEIN (FRAGMENT).
 GN DKFZP586B147.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-UTERUS;
 RA Koehler K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL117448; CAB55932.1; -.
 DR InterPro: IPR001194; DENN.
 DR InterPro: IPR001024; LH2.
 DR InterPro: IPR004012; Run.
 DR InterPro: IPR001194; DENN.
 DR Pfam: PF02141; DENN; 1.
 DR Pfam: PF01477; PLAT; 1.
 DR Pfam: PF02759; RUN; 1.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 1291 AA; 147377 MW; E82F3F38EAD3BC42 CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
 |||||
 Db 238 OPPPLP 243

Search completed: August 8, 2002, 07:49:13

Job time: 146 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2002, 07:46:02 ; Search time 12.95 Seconds
(without alignments)
13.203 Million cell updates/sec

Title: US-09-641-801-1
Perfect score: 42
Sequence: 1 MQPPPLP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :
1: /cgnt2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgnt2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgnt2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgnt2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgnt2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgnt2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	88.1	278	4	US-08-339-214-16
2	37	88.1	278	4	US-08-339-214-16
3	37	88.1	667	2	US-08-718-661-2
4	36	85.7	349	1	US-08-118-270-71
5	36	85.7	349	5	PCR-US93-08528-71
6	35	83.3	23	1	US-08-268-251-34
7	35	83.3	23	5	PCR-US93-01112-34
8	34	81.0	194	3	US-08-822-264-4
9	34	81.0	195	2	US-08-822-264-4
10	34	81.0	223	3	US-08-822-264-3
11	34	81.0	376	4	US-09-056-556-202
12	34	81.0	409	4	US-09-029-755C-2
13	34	81.0	608	4	US-09-413-814-92
14	34	81.0	1182	4	US-09-293-505-7
15	34	81.0	1213	4	US-09-413-814-79
16	33	78.6	10	4	US-08-602-999A-278
17	33	78.6	95	4	US-09-314-268-132
18	33	78.6	324	4	US-08-732-749-4
19	33	78.6	400	2	US-08-747-887-2
20	33	78.6	726	4	US-09-126-980-2
21	33	78.6	726	4	US-09-126-980-2
22	33	78.6	834	3	US-08-539-205A-6
23	33	78.6	902	1	US-08-396-479B-6
24	33	78.6	902	1	US-08-818-823-6
25	33	78.6	1203	4	US-09-207-857-2
26	33	78.6	1203	4	US-09-293-505-2
27	33	78.6	1784	4	US-09-040-738-2

28	33	78.6	1784	4	US-08-652-426A-2	Sequence 2, App11
29	32	76.2	7	5	US-08-230-047-40	Sequence 40, App1
30	32	76.2	7	5	PCR-US94-01840-11	Sequence 11, App1
31	32	76.2	8	2	US-08-612-857-8	Sequence 8, App1
32	32	76.2	10	1	US-08-230-047-12	Sequence 12, App1
33	32	76.2	10	1	US-08-230-047-14	Sequence 14, App1
34	32	76.2	10	1	US-08-212-190A-8	Sequence 8, App11
35	32	76.2	10	2	US-08-900-321-8	Sequence 8, App11
36	32	76.2	10	3	PCR-US95-03610-8	Sequence 8, App11
37	32	76.2	10	5	US-08-899-595-8	Sequence 8, App11
38	32	76.2	11	1	US-08-336-343A-25	Sequence 25, App1
39	32	76.2	11	4	US-08-652-877-22	Sequence 22, App1
40	32	76.2	11	4	US-08-652-877-32	Sequence 32, App1
41	32	76.2	11	4	US-08-476-515A-22	Sequence 22, App1
42	32	76.2	11	4	US-08-476-515A-32	Sequence 32, App1
43	32	76.2	12	1	US-08-230-047-17	Sequence 17, App1
44	32	76.2	12	4	US-08-602-999A-251	Sequence 251, App
45	32	76.2	12	4	US-08-602-999A-267	Sequence 267, App
46	32	76.2	14	1	US-08-185-432-12	Sequence 12, App1
47	32	76.2	14	3	US-08-630-916A-63	Sequence 63, App1
48	32	76.2	14	4	US-08-630-915A-186	Sequence 186, App
49	32	76.2	15	1	US-08-185-432-10	Sequence 10, App1
50	32	76.2	15	2	US-08-769-745-31	Sequence 31, App1
51	32	76.2	15	4	US-08-602-999A-356	Sequence 356, App
52	32	76.2	15	4	US-08-602-999A-385	Sequence 385, App
53	32	76.2	15	4	US-08-602-999A-404	Sequence 404, App
54	32	76.2	15	4	US-08-602-999A-414	Sequence 414, App
55	32	76.2	15	4	US-08-602-999A-435	Sequence 435, App
56	32	76.2	15	4	US-08-602-999A-436	Sequence 436, App
57	32	76.2	15	4	US-08-602-999A-441	Sequence 441, App
58	32	76.2	15	4	US-08-602-999A-452	Sequence 452, App
59	32	76.2	16	1	US-08-185-432-9	Sequence 9, App11
60	32	76.2	16	1	US-08-185-432-14	Sequence 14, App1
61	32	76.2	16	4	US-08-602-999A-197	Sequence 197, App
62	32	76.2	17	1	US-08-230-047-10	Sequence 10, App1
63	32	76.2	17	4	US-08-602-999A-360	Sequence 360, App
64	32	76.2	17	4	US-08-602-999A-374	Sequence 374, App
65	32	76.2	18	4	US-08-602-999A-320	Sequence 320, App
66	32	76.2	18	4	US-08-602-999A-409	Sequence 409, App
67	32	76.2	19	4	US-08-630-915A-169	Sequence 169, App
68	32	76.2	20	4	US-08-602-999A-358	Sequence 358, App
69	32	76.2	21	2	US-08-459-568-59	Sequence 59, App1
70	32	76.2	21	2	US-08-399-411-59	Sequence 59, App1
71	32	76.2	21	3	US-08-516-859A-59	Sequence 59, App1
72	32	76.2	22	2	US-08-598-873-28	Sequence 28, App1
73	32	76.2	22	4	US-08-605-430-28	Sequence 28, App1
74	32	76.2	22	4	US-08-605-430-28	Sequence 28, App1
75	32	76.2	25	1	US-08-185-432-13	Sequence 13, App1
76	32	76.2	25	2	US-08-383-753-21	Sequence 21, App1
77	32	76.2	25	2	US-08-386-772-21	Sequence 21, App1
78	32	76.2	25	4	US-08-559-512-21	Sequence 21, App1
79	32	76.2	28	1	US-09-512-983-21	Sequence 21, App1
80	32	76.2	28	2	US-08-732-751-16	Sequence 16, App1
81	32	76.2	28	2	US-08-459-568-58	Sequence 58, App1
82	32	76.2	28	2	US-08-399-411-58	Sequence 58, App1
83	32	76.2	28	2	US-08-769-745-39	Sequence 39, App1
84	32	76.2	28	3	US-08-516-859A-58	Sequence 58, App1
85	32	76.2	29	1	US-08-230-047-9	Sequence 9, App11
86	32	76.2	32	1	US-08-732-751-14	Sequence 14, App1
87	32	76.2	57	2	US-08-318-837-50	Sequence 50, App1
88	32	76.2	140	3	US-08-850-961-12	Sequence 12, App1
89	32	76.2	140	3	US-09-479-776-12	Sequence 12, App1
90	32	76.2	141	3	US-08-850-961-10	Sequence 10, App1
91	32	76.2	141	4	US-09-479-776-10	Sequence 10, App1
92	32	76.2	157	1	US-09-267-177-19	Sequence 19, App1
93	32	76.2	181	4	US-08-185-432-6	Sequence 6, App11
94	32	76.2	220	3	US-08-822-264-1	Sequence 1, App11
95	32	76.2	255	2	US-08-776-156-2	Sequence 2, App11
96	32	76.2	256	3	US-08-659-188-2	Sequence 2, App11
97	32	76.2	256	3	US-08-555-221-2	Sequence 2, App11
98	32	76.2	256	3	US-08-655-221-2	Sequence 2, App11
99	32	76.2	281	4	US-09-398-326-2	Sequence 2, App11
100	32	76.2	281	3	US-08-810-453-2	Sequence 2, App11
					US-08-815-190A-2	Sequence 2, App11

101	32	76.2	281	4	US-09-290-640-25	Sequence 25, Appl
102	32	76.2	281	4	US-09-479-524-3	Sequence 3, Appl
103	32	76.2	281	4	US-08-339-214-8	Sequence 8, Appl
104	32	76.2	281	4	US-08-339-214-30	Sequence 30, Appl
105	32	76.2	281	5	PCT-US95-00362-2	Sequence 2, Appl
106	32	76.2	287	4	US-09-031-962D-2	Sequence 2, Appl
107	32	76.2	288	4	US-08-545-196B-21	Sequence 21, Appl
108	32	76.2	288	4	US-08-028-327-3	Sequence 2, Appl
109	32	76.2	300	2	US-08-592-214A-2	Sequence 2, Appl
110	32	76.2	300	3	US-09-149-976-2	Sequence 2, Appl
111	32	76.2	311	2	US-08-318-837-7	Sequence 7, Appl
112	32	76.2	311	4	US-09-179-558-66	Sequence 66, Appl
113	32	76.2	324	2	US-08-484-993B-22	Sequence 22, Appl
114	32	76.2	324	2	US-08-484-158B-22	Sequence 22, Appl
115	32	76.2	324	2	US-08-484-596A-22	Sequence 22, Appl
116	32	76.2	324	2	US-08-480-150A-22	Sequence 22, Appl
117	32	76.2	324	3	US-08-458-731-22	Sequence 22, Appl
118	32	76.2	324	3	US-08-149-223A-22	Sequence 22, Appl
119	32	76.2	338	1	US-08-218-686-2	Sequence 2, Appl
120	32	76.2	338	3	US-08-460-242-2	Sequence 2, Appl
121	32	76.2	384	2	US-08-872-302-4	Sequence 4, Appl
122	32	76.2	393	2	US-09-026-587-3	Sequence 3, Appl
123	32	76.2	393	2	US-09-227-420-3	Sequence 3, Appl
124	32	76.2	403	2	US-08-592-383-4	Sequence 4, Appl
125	32	76.2	413	3	US-08-836-582-2	Sequence 2, Appl
126	32	76.2	413	4	US-09-265-566-2	Sequence 4, Appl
127	32	76.2	416	4	US-08-764-870-4	Sequence 4, Appl
128	32	76.2	416	4	US-08-980-115-4	Sequence 4, Appl
129	32	76.2	418	4	US-09-026-587-1	Sequence 1, Appl
130	32	76.2	421	4	US-09-362-473-14	Sequence 14, Appl
131	32	76.2	421	4	US-09-362-473-14	Sequence 14, Appl
132	32	76.2	433	1	US-08-615-170-19	Sequence 19, Appl
133	32	76.2	433	1	US-08-615-170-19	Sequence 19, Appl
134	32	76.2	434	4	US-09-252-292C-26	Sequence 26, Appl
135	32	76.2	445	1	US-08-900-148-2	Sequence 2, Appl
136	32	76.2	453	1	US-08-417-330A-18	Sequence 18, Appl
137	32	76.2	462	2	US-08-592-383-2	Sequence 2, Appl
138	32	76.2	462	2	US-08-095-728B-4	Sequence 4, Appl
139	32	76.2	462	5	PCT-US92-02320A-4	Sequence 4, Appl
140	32	76.2	462	5	PCT-US92-02320A-4	Sequence 4, Appl
141	32	76.2	462	6 <th>5171671-2</th> <th>Patent No. 5171671</th>	5171671-2	Patent No. 5171671
142	32	76.2	480	1	US-07-882-292-2	Sequence 2, Appl
143	32	76.2	480	2 <th>US-08-331-644-4</th> <th>Sequence 2, Appl</th>	US-08-331-644-4	Sequence 2, Appl
144	32	76.2	480	5 <th>PCT-US93-04102-2</th> <th>Sequence 2, Appl</th>	PCT-US93-04102-2	Sequence 2, Appl
145	32	76.2	488	1 <th>US-07-794-393-2</th> <th>Sequence 2, Appl</th>	US-07-794-393-2	Sequence 2, Appl
146	32	76.2	488	1 <th>US-08-001-711-2</th> <th>Sequence 2, Appl</th>	US-08-001-711-2	Sequence 2, Appl
147	32	76.2	488	3 <th>US-08-704-711A-22</th> <th>Sequence 22, Appl</th>	US-08-704-711A-22	Sequence 22, Appl
148	32	76.2	492	4 <th>US-08-448-489-11</th> <th>Sequence 11, Appl</th>	US-08-448-489-11	Sequence 11, Appl
149	32	76.2	492	1 <th>US-08-469-486-2</th> <th>Sequence 2, Appl</th>	US-08-469-486-2	Sequence 2, Appl
150	32	76.2	492	2 <th>US-08-469-658-2</th> <th>Sequence 2, Appl</th>	US-08-469-658-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-339-214-16
Sequence 16, Application US/08339214
Patent No. 6348334
GENERAL INFORMATION:
APPLICANT: Nagata, Shigikazu
APPLICANT: Suda, Takashi
APPLICANT: Takahashi, Tomoniro
APPLICANT: Nakamura, No. 6348334io
TITLE OF INVENTION: A Fas Ligand, A Fragment Thereof and DNA
TITLE OF INVENTION: Encoding the Same
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA

ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/339,214
FILING DATE: 10-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1110-139P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 278 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-214-16

Query Match 88.1% Score 37; DB 4; Length 278;
Best Local Similarity 100.0%; Pred. No. 1,1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPLP 7
Db 60 QPPLP 65

RESULT 2
US-08-339-214-26
Sequence 26, Application US/08339214
Patent No. 6348334
GENERAL INFORMATION:
APPLICANT: Nagata, Shigikazu
APPLICANT: Suda, Takashi
APPLICANT: Takahashi, Tomoniro
APPLICANT: Nakamura, No. 6348334io
TITLE OF INVENTION: A Fas Ligand, A Fragment Thereof and DNA
TITLE OF INVENTION: Encoding the Same
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/339,214
FILING DATE: 10-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1110-139P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 278 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-214-26

Query Match 88.1%; Score 37; DB 4; Length 278;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
DB 60 QPPPLP 65

RESULT 3
US-08-718-661-2
Sequence 2, Application US/08718661
Patent No. 5876972

GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Nucleic acid molecules coding for mammalian
tumor suppressor proteins and methods for their isolation
TITLE OF INVENTION: tumor suppressor proteins and methods for their isolation
NUMBER OF SEQUENCES: 15
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718,661
FILING DATE:
CLASSIFICATION: 530
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 667 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-718-661-2

Query Match 88.1%; Score 37; DB 2; Length 667;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
DB 527 QPPPLP 532

RESULT 4
US-08-118-270-71
Sequence 71, Application US/08118270
Patent No. 5508384
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
APPLICANT: Schuster, David I.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-118-270-71

Query Match 85.7%; Score 36; DB 1; Length 349;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOPPLP 7
DB 290 MOPPLP 296

RESULT 5
PCT-US93-08528-71
Sequence 71, Application PC/TUS9308528
GENERAL INFORMATION:
APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
PCT-US93-08528-71

Query Match 85.7%; Score 36; DB 5; Length 349;
Best Local Similarity 85.7%; Pred. No. 1.de+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
DB 290 MOPPPK 296

RESULT 6
US-08-268-251-34
Sequence 34, Application US/08268251
Patent No. 5583475
GENERAL INFORMATION:
APPLICANT: Jamieson, Gordon A
APPLICANT: Dedman, John R
APPLICANT: Kaetzel, Marcia A
TITLE OF INVENTION: Calmodulin-binding Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/268, 251
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831, 219
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D
REGISTRATION NUMBER: 31, 259
REFERENCE/DOCKET NUMBER: 272, 001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-268-251-34

Query Match 83.3%; Score 35; DB 1; Length 23;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
DB 16 MOPPPP 22

RESULT 7
PCT-US93-01112-34
Sequence 34, Application PC/TUS9301112
GENERAL INFORMATION:
APPLICANT: Jamieson, Gordon A

APPLICANT: Dedman, John R
APPLICANT: Kaetzel, Marcia A
TITLE OF INVENTION: Calmodulin-binding peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01112
FILING DATE: 19930208
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831, 219
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D
REGISTRATION NUMBER: 31, 259
REFERENCE/DOCKET NUMBER: 272, 001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-01112-34

Query Match 83.3%; Score 35; DB 5; Length 23;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
DB 16 MOPPPP 22

RESULT 8
US-08-822-264-4
Sequence 4, Application US/08822264
Patent No. 6033869
GENERAL INFORMATION:
APPLICANT: Goll, Surya K.
APPLICANT: Hillman, Jennifer L.
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: NOVEL HUMAN CYTOKINE/STEROID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/822,264
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1657409
US-08-822-264-4

Query Match 81.0%; Score 34; DB 3; Length 194;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
:|||||
DB 60 EPPPLP 65

RESULT 9
US-08-960-022-6
Sequence 6, Application US/08960022
Patent No. 5976837.
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Racle, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Yikhi
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/960,022
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,333
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 195 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-960-022-6

Query Match 81.0%; Score 34; DB 2; Length 195;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
:|||||
DB 61 EPPPLP 66

RESULT 10
US-08-822-264-3
Sequence 3, Application US/08822264
Patent No. 603869
GENERAL INFORMATION:
APPLICANT: Goli, Surya K.
APPLICANT: Hillman, Jennifer L.
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: NOVEL HUMAN CYTOKINE/STEROID
TITLE OF INVENTION: RECEPTOR PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,264
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 158818
US-08-822-264-3

Query Match 81.0%; Score 34; DB 3; Length 223;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 OPPPLP 7
:|||||
DB 61 EPPPLP 66

RESULT 11

US-09-056-556-202
Sequence 202, Application US/09056556
Patent No. 6350456
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
FILING DATE: 07-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 202:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-056-556-202

Query Match 81.0%; Score 34; DB 4; Length 376;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 OPPPLP 7
:|||||
DB 212 OPPPV 217

RESULT 12

US-09-029-755C-2
Sequence 2, Application US/09029755C
Patent No. 6326477
GENERAL INFORMATION:
APPLICANT: ILMEN, Marja
APPLICANT: SOEDERLUND, Hans
APPLICANT: PENTTILA, Merja
TITLE OF INVENTION: PROCESS FOR MODIFYING GLUCOSE REPRESSION
FILE REFERENCE: Substitute Sequence listing-09-029755
Patent No. 6326477
CURRENT APPLICATION NUMBER: US/09/029,755C
CURRENT FILING DATE: 1998-03-02
PRIOR APPLICATION NUMBER: PCT/FI96/00463
PRIOR FILING DATE: 1996-08-30
PRIOR APPLICATION NUMBER: FI 954123
PRIOR FILING DATE: 1995-09-01

NUMBER OF SEQ ID NOS: 21
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 409
TYPE: PRT
ORGANISM: Trichocheima harzianum T3
US-09-029-755C-2

Query Match 81.0%; Score 34; DB 4; Length 409;
Best Local Similarity 71.4%; Pred. No. 4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOPPLP 7
:|||||
DB 148 MPPPPV 154

RESULT 13

US-09-413-814-92
Sequence 92, Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
FILE REFERENCE: PCT/US 99/23535
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 92
LENGTH: 608
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-92

Query Match 81.0%; Score 34; DB 4; Length 608;
Best Local Similarity 83.3%; Pred. No. 5.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 OPPPLP 7
:|||||
DB 437 EPPPLP 442

RESULT 14

US-09-293-505-7
Sequence 7, Application US/09293505
Patent No. 6348575
GENERAL INFORMATION:
APPLICANT: de Sauvage, Frederic
APPLICANT: Carpenter, David A.
TITLE OF INVENTION: Patched-2
FILE REFERENCE: P1405R1
CURRENT APPLICATION NUMBER: US/09/293,505
CURRENT FILING DATE: 1999-04-15
EARLIER APPLICATION NUMBER: US 60/081,884
EARLIER FILING DATE: 1998-04-15
NUMBER OF SEQ ID NOS: 32

SEQ ID NO 7
LENGTH: 1182
TYPE: PRT
ORGANISM: Mus musculus
US-09-293-505-7

Query Match 81.0%; Score 34; DB 4; Length 1182;
Best Local Similarity 71.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOPPLP 7
: |||||
DB 1164 LPPPLP 1170

RESULT 15
US-09-413-814-79
Sequence 79, Application US/09413814
Patent No. 6225064

GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bloeker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hoffe, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: heteropolypeptide compounds
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 79
LENGTH: 1213
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-79

Query Match 81.0%; Score 34; DB 4; Length 1213;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
: |||||
DB 1041 EPPPLP 1046

RESULT 16
US-08-602-999A-278
Sequence 278, Application US/08602999A
Patent No. 6184205

GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: KAY, Brian K.
APPLICANT: THORN, Judith M.
APPLICANT: QUILIAM, Lawrence A.
APPLICANT: DER, Channing J.
APPLICANT: FOWLES, Dana M.
APPLICANT: RIDER, James E.
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
ISOLATING AND USING SAME
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,999A
FILING DATE: 16-FEB-1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 278:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-602-999A-278

Query Match 78.6%; Score 33; DB 4; Length 10;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
: |||||
DB 1 KPPPLP 6

RESULT 17
US-09-314-268-132
Sequence 132, Application US/09314268
Patent No. 6346377

GENERAL INFORMATION:
APPLICANT: Doordat, John
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
FILE REFERENCE: 3789/80902
CURRENT APPLICATION NUMBER: US/09/314,268
CURRENT FILING DATE: 1999-03-19
EARLIER APPLICATION NUMBER: 09/314,268
EARLIER FILING DATE: 1999-05-18
NUMBER OF SEQ ID NOS: 179
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 132
LENGTH: 95
TYPE: PRT
ORGANISM: Human papillomavirus type 25
US-09-314-268-132

Query Match 78.6%; Score 33; DB 4; Length 95;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOPPLP 7
: |||||
DB 80 LPPPLP 86

RESULT 18

US-08-732-749-4
Sequence 4, Application US/08732749
Patent No. 6300062
GENERAL INFORMATION:
APPLICANT: CERNY, Radim
APPLICANT: SLAVY, Ivan
APPLICANT: HAMARSTROEM, Lars
APPLICANT: WURZ, Tilmann
APPLICANT: FONG, Cheng D.
TITLE OF INVENTION: ENAMEL MATRIX RELATED POLYPEPTIDE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/732,749
FILING DATE: 18-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/005,634
FILING DATE: 19-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: CERNY-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 324 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-732-749-4

Query Match 78.6%; Score 33; DB 4; Length 324;
Best Local Similarity 71.4%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOPPLP 7
DB 12 VHPPLP 18

RESULT 19
US-08-747-887-2
Sequence 2, Application US/08747887
Patent No. 5853734
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: MOORE, Patrick S.
TITLE OF INVENTION: Glycoprotein L And Glycoprotein M From Kaposi's
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA Encoding
TITLE OF INVENTION: Same And Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York

STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/747,887
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 52429
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-747-887-2

Query Match 78.6%; Score 33; DB 2; Length 400;
Best Local Similarity 83.3%; Pred. No. 5.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
DB 356 KPPPLP 361

RESULT 20
US-09-126-980-2
Sequence 2, Application US/09126980
Patent No. 6270936
GENERAL INFORMATION:
APPLICANT: Jones, Katherine
APPLICANT: Wei, Ping
APPLICANT: Garber, Mitchell
APPLICANT: Fang, Shi-Min
TITLE OF INVENTION: A TRANSCRIPTIONAL COACTIVATOR THAT
TITLE OF INVENTION: INTERACTS WITH TAT PROTEIN AND REGULATES ITS
TITLE OF INVENTION: BINDING TO TAR RNA, METHODS FOR MODULATING TAT
TITLE OF INVENTION: TRANSCRIPTION, AND USES THEREFOR
FILE REFERENCE: SALK2231
CURRENT APPLICATION NUMBER: US/09/126,980
CURRENT FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/069,341
EARLIER FILING DATE: 1997-12-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 726
TYPE: PRT
ORGANISM: Homo sapiens
US-09-126-980-2

Query Match 78.6%; Score 33; DB 4; Length 726;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QPPPLP 7
DB 709 RPPPLP 714

RESULT 21
US-09-476-482-2
Sequence 2, Application US/09476482
Patent No. 6284456
GENERAL INFORMATION:
APPLICANT: Jones, Katherine A.
APPLICANT: Wei, Ping
APPLICANT: Garber, Mitchell
APPLICANT: Fang, Shi-Min
TITLE OF INVENTION: A TRANSCRIPTIONAL COACTIVATOR THAT
TITLE OF INVENTION: INTERACTS WITH TAT PROTEIN AND REGULATES ITS BINDING TO TAR
TITLE OF INVENTION: RNA, METHODS FOR MODULATING TAT TRANSCRIPTION, AND USES
FILE REFERENCE: SALK2230-2
CURRENT APPLICATION NUMBER: US/09/476,482
CURRENT FILING DATE: 1999-12-30
EARLIER APPLICATION NUMBER: 09/126,980
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 726
TYPE: PRT
ORGANISM: Homo sapiens
US-09-476-482-2

Query Match 78.6%; Score 33; DB 4; Length 726;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
:|||||
Db 709 RPPPLP 714

RESULT 22
US-08-539-205A-6
Sequence 6, Application US/08539205A
Patent No. 6001619
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Calligaris, Maureen
APPLICANT: Nefsky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/539,205A
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: C5V-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-539-205A-6

Query Match 78.6%; Score 33; DB 3; Length 834;
Best Local Similarity 71.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOPPLP 7
:|||||
Db 69 LPPPLP 75

RESULT 23
US-08-396-479B-6
Sequence 6, Application US/08396479B
Patent No. 5612455
GENERAL INFORMATION:
APPLICANT: HOEY, Timothy
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/396,479B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59450-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 210 277299

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 902 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-396-479B-6

Query Match 78.6%; Score 33; DB 1; Length 902;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 OPPPLP 7
:|||||
Db 812 RPPPLP 817

RESULT 24
US-08-818-823-6
Sequence 6, Application US/08818823
Patent No. 5708158
GENERAL INFORMATION:
APPLICANT: HOEY, Timothy
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818, 823
FILING DATE: 14-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/396,479
FILING DATE: 02-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59450-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8771
TELEFAX: (415) 494-8700
TELEX: 210 277299
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 902 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-818-823-6

Query Match 78.6%; Score 33; DB 1; Length 902;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MOPPLP 7
:|||||
DB 812 RPPPLP 817

RESULT 25
US-09-207-857-2
Sequence 2, Application US/09207857
Patent No. 6309879
GENERAL INFORMATION:
APPLICANT: Buncroft, David A.
TITLE OF INVENTION: HUMAN PATCHED GENES AND PROTEINS, AND USES RELATED
TITLE OF INVENTION: THEREO
FILE REFERENCE: ONV-05001
CURRENT APPLICATION NUMBER: US/09/207,857
CURRENT FILING DATE: 1998-12-08
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1203
TYPE: PRT
ORGANISM: human
US-09-207-857-2

Query Match 78.6%; Score 33; DB 4; Length 1203;
Best Local Similarity 71.4%; Pred. No. 1.5e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOPPLP 7
:|||||
DB 1164 IHPPPLP 1170

RESULT 26

US-09-293-505-2
Sequence 2, Application US/09293505
Patent No. 6348575
GENERAL INFORMATION:
APPLICANT: de Sauvage, Frederic
APPLICANT: Carpenter, David A.
TITLE OF INVENTION: Patched-2
FILE REFERENCE: P1405R1
CURRENT APPLICATION NUMBER: US/09/293,505
CURRENT FILING DATE: 1999-04-15
EARLIER APPLICATION NUMBER: US 60/081,884
EARLIER FILING DATE: 1998-04-15
NUMBER OF SEQ ID NOS: 32
SEQ ID NO 2
LENGTH: 1203
TYPE: PRT
ORGANISM: Homo sapiens
US-09-293-505-2

Query Match 78.6%; Score 33; DB 4; Length 1203;
Best Local Similarity 71.4%; Pred. No. 1.5e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOPPLP 7
:|||||
DB 1164 IHPPPLP 1170

RESULT 27
US-09-040-738-2
Sequence 2, Application US/09040738
Patent No. 6207374
GENERAL INFORMATION:
APPLICANT: Sampson et al.
TITLE OF INVENTION: Tuberosus sclerosis 2 Gene and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathleen M. Williams, Banner & Witcoff,
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,738
FILING DATE: Concurrently herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9326470.3
FILING DATE: 24-December-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9411900.5
FILING DATE: 14-June-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB94/02823
FILING DATE: 23-December-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/652,426
FILING DATE: 30-May-1996
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3265/73963
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100

TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1784 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-040-738-2

Query Match 78.6%; Score 33; DB 4; Length 1784;
Best Local Similarity 83.3%; Pred. No. 2.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 PPPLP 7
DB 1262 KPPPLP 1267

RESULT 28
US-08-652-426A-2
Sequence 2, Application US/08652426A
Patent No. 6232452
GENERAL INFORMATION:
APPLICANT: Sampson et al.
TITLE OF INVENTION: Tubercous Sclerosis 2 Gene and Uses
TITLE OF INVENTION: Tubercous Sclerosis 2 Gene and Uses
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Kathleen M. Williams, Banner & Wilcoff,
ADDRESSEE: Ltd.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,426A
FILING DATE: 01-October-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9326470.3
FILING DATE: 12/24/93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9411900.5
FILING DATE: 06/14/94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB94/02823
FILING DATE: 12/23/94
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 96,738
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1784 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-426A-2

Query Match 78.6%; Score 33; DB 4; Length 1784;
Best Local Similarity 83.3%; Pred. No. 2.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 PPPLP 7
DB 1262 KPPPLP 1267

RESULT 29
US-08-230-047-40
Sequence 40, Application US/08230047
Patent No. 5541109
GENERAL INFORMATION:
APPLICANT: Seafoss III, George H.
APPLICANT: Ivashchenko, Yuri D.
APPLICANT: Jaye, Michael C.
TITLE OF INVENTION: EXPRESSION CLONING OF C-SRC SH3 BINDING
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: word 5.0 (patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/230,047
FILING DATE: 19-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Rosanne
REGISTRATION NUMBER: 32,534
REFERENCE/DOCKET NUMBER: A1465-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3817
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-230-047-40

Query Match 76.2%; Score 32; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPPLP 7
DB 1 PPLP 5

RESULT 30
PCT-US94-01840-11
Sequence 11, Application PC/TUS9401840
GENERAL INFORMATION:
APPLICANT: Christopher E. Rudd
APPLICANT: Prasad Kanteti
APPLICANT: Lewis Cantley
TITLE OF INVENTION: CD4 MEDIATED MODULATION OF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts

COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01840
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,915
FILING DATE: February 26, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Janis K. Fraser
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/063001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 7
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: Linear
PCT-US94-01840-11

Query Match 76.2%; Score 32; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PPLP 7
11111
Db 1 PPLP 5

RESULT 31
US-08-612-857-8
Sequence 8, Application US/08612857
Patent No. 5831048
GENERAL INFORMATION:
APPLICANT: SCHWEIGHOFFER, Fabien
APPLICANT: TOCQUE, Bruno
TITLE OF INVENTION: GRB3-3 GENE, VARIANTS AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd. 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,857
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93-10971
FILING DATE: 15-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR94/00542
FILING DATE: 09-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.

REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: ST93044-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)454-3839
TELEFAX: (610)454-3808
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..8
OTHER INFORMATION: /note="3BPI peptide"

Query Match 76.2%; Score 32; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PPLP 7
11111
Db 1 PPLP 5

RESULT 32
US-08-230-047-12
Sequence 12, Application US/08230047
Patent No. 5541109
GENERAL INFORMATION:
APPLICANT: Seartoss III, George H.
APPLICANT: Ivashchenko, Yuri D.
TITLE OF INVENTION: EXPRESSION CLONING OF C-SRC SH3 BINDING
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: word 5.0 (Patentlin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/230,047
FILING DATE: 19-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Rosanne
REGISTRATION NUMBER: 32,534
REFERENCE/DOCKET NUMBER: A1465-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3817
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-230-047-12

Query Match 76.2%; Score 32; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPLP 7
|||||
DB 5 PPLP 9

RESULT 33

US-08-230-047-14
; Sequence 14, Application US/08230047
; Patent No. 5541109
; GENERAL INFORMATION:
; APPLICANT: Seartoss III, George H.
; APPLICANT: Ivashchenko, Yuri D.
; APPLICANT: Jaye, Michael C.
; TITLE OF INVENTION: EXPRESSION CLONING OF C-SRC SH3 BINDING
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.0 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/230,047
; FILING DATE: 19-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne
; REGISTRATION NUMBER: 32,534
; REFERENCE/DOCKET NUMBER: A1465-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3817
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-230-047-14

Query Match 76.2%; Score 32; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPLP 7
|||||
DB 6 PPLP 10

RESULT 34

US-08-212-190A-8
; Sequence 8, Application US/08212190A
; Patent No. 5652223
; GENERAL INFORMATION:
; APPLICANT: KOHN, Elise C.
; APPLICANT: LIOTTA, Lance A.
; APPLICANT: KIM, Young Sook
; TITLE OF INVENTION: DNA ENCODING CAI RESISTANCE PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew

STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/212,190A
FILING DATE: 14-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 15280-204US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-212-190A-8

Query Match 76.2%; Score 32; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPLP 7
|||||
DB 5 PPLP 9

RESULT 35

US-08-900-321-8
; Sequence 8, Application US/08900321
; Patent No. 5981712
; GENERAL INFORMATION:
; APPLICANT: Kohn, Elise C.
; APPLICANT: Liotta, Lance A.
; APPLICANT: Kim, Young S.
; TITLE OF INVENTION: DNA Encoding CAI Resistance Proteins and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,321
; FILING DATE: 25-JUL-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,190
; FILING DATE: 14-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom

REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-204100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-900-321-8

Query Match 76.2%; Score 32; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PPLP 7
DB 5 PPLP 9

RESULT 36
US-08-899-595-8
Sequence 8, Application US/08895955
Patent No. 6111072
GENERAL INFORMATION:
APPLICANT: Natumiya, Shuh
APPLICANT: Takahashi, No. 6111072uak1
TITLE OF INVENTION: RHO TARGET PROTEIN HUMAN MDIA AND GENE
TITLE OF INVENTION: ENCODING SAME
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,595
FILING DATE: 24-JUL-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-242701
FILING DATE: 26-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 9-90170
FILING DATE: 25-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Stephen A. Bent
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 049441/0112
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site

LOCATION: 10
OTHER INFORMATION: /Product="G/S/A/V"
US-08-899-595-8

Query Match 76.2%; Score 32; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PPLP 7
DB 4 PPLP 8

RESULT 37
PCT-US95-03610-8
Sequence 8, Application PC/TUS9503610.
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA ENCODING CAI RESISTANCE PROTEINS AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 10
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/03610
FILING DATE: 14-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,190
FILING DATE: 14-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-204000PC
REFERENCE/DOCKET NUMBER: DHS Ref. No. E-112-94/0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-03610-8

Query Match 76.2%; Score 32; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PPLP 7
DB 5 PPLP 9

RESULT 38
US-08-336-343A-25
Sequence 25, Application US/08336343A
Patent No. 5677144
GENERAL INFORMATION:
APPLICANT: Ullrich, Axel
APPLICANT: Alves, Frauke
TITLE OF INVENTION: CCK-2, A NO. 5677144e1 Receptor Tyrosine Kinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,343A
FILING DATE: 08-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-065
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-336-343A-25

Query Match 76.2%; Score 32; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPLP 7
11111
DB 4 PPLP 8

RESULT 39

US-08-652-877-22
Sequence 22, Application US/08652877
Patent No. 6187548
GENERAL INFORMATION:
APPLICANT: Akersstrom, Goran
APPLICANT: Juhlin, Claes
APPLICANT: Rask, Lars
APPLICANT: Crumley, Gregg R.
APPLICANT: Morse, Clarence C.
APPLICANT: Murray, Edward M.
APPLICANT: Hjaln, Goran
TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
TITLE OF INVENTION: Thereof and DNA Encoding Same
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcoia Rd., 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426-0107
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.5.1
SOFTWARE: Word 6.0 (patentln)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,877
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15203
FILING DATE: 22-NOV-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,836
FILING DATE: 23-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,314
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky, Martin
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: A1355E-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-454-3816
TELEFAX: 610-454-3808
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: Internal
US-08-652-877-22

Query Match 76.2%; Score 32; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPLP 7
11111
DB 4 PPLP 8

RESULT 40

US-08-652-877-32
Sequence 32, Application US/08652877
Patent No. 6187548
GENERAL INFORMATION:
APPLICANT: Akersstrom, Goran
APPLICANT: Juhlin, Claes
APPLICANT: Rask, Lars
APPLICANT: Crumley, Gregg R.
APPLICANT: Morse, Clarence C.
APPLICANT: Murray, Edward M.
APPLICANT: Hjaln, Goran
TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
TITLE OF INVENTION: Thereof and DNA Encoding Same
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcoia Rd., 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426-0107
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.5.1
SOFTWARE: Word 6.0 (patentln)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,877
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15203
FILING DATE: 22-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,836
FILING DATE: 23-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,314

FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky, Martin
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: A1355E-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-454-3816
TELEFAX: 610-454-3808
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS:
MOLECULE TYPE: linear
TOPOLOGY: linear
HYPOTHETICAL: NO
FRAGMENT TYPE: Internal
US-08-652-877-32

Query Match 76.2%; Score 32; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPLP 7
|||||
DB 4 PPLP 8

RESULT 41
US-08-476-515A-22
Sequence 22, Application US/08476515A
Patent No. 6239270
GENERAL INFORMATION:
APPLICANT: Akertstrom, Goran
APPLICANT: Junlin, Claes
APPLICANT: Raak, Lars
APPLICANT: Crumley, Gregg R.
APPLICANT: Morse, Clarence C.
APPLICANT: Murray, Edward M.
APPLICANT: Hjaln, Goran
TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
TITLE OF INVENTION: Thereof and DNA Encoding Same
NUMBER OF SEQUENCES: 84
CORRESPONDENCE ADDRESS:
ADDRESSEE: Martin Savitzky
STREET: Rhone-Poulenc Rorer Inc., 500 Arcola Rd.,
STREET: 3C43,
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426-0107
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Compaq PC
OPERATING SYSTEM: Windows 95
SOFTWARE: Word 7.0 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,515A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,836
FILING DATE: 23-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/SE94/00483
FILING DATE: 24-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9301764-8
FILING DATE: 24-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky, Martin
REGISTRATION NUMBER: 29,699

REFERENCE/DOCKET NUMBER: A1355D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-454-3816
TELEFAX: 610-454-3808
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS:
MOLECULE TYPE: linear
TOPOLOGY: linear
HYPOTHETICAL: NO
FRAGMENT TYPE: Internal
US-08-476-515A-22

Query Match 76.2%; Score 32; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPLP 7
|||||
DB 4 PPLP 8

RESULT 42
US-08-476-515A-32
Sequence 32, Application US/08476515A
Patent No. 6239270
GENERAL INFORMATION:
APPLICANT: Akertstrom, Goran
APPLICANT: Junlin, Claes
APPLICANT: Raak, Lars
APPLICANT: Crumley, Gregg R.
APPLICANT: Morse, Clarence C.
APPLICANT: Murray, Edward M.
APPLICANT: Hjaln, Goran
TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
TITLE OF INVENTION: Thereof and DNA Encoding Same
NUMBER OF SEQUENCES: 84
CORRESPONDENCE ADDRESS:
ADDRESSEE: Martin Savitzky
STREET: Rhone-Poulenc Rorer Inc., 500 Arcola Rd.,
STREET: 3C43,
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426-0107
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Compaq PC
OPERATING SYSTEM: Windows 95
SOFTWARE: Word 7.0 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,515A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,836
FILING DATE: 23-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/SE94/00483
FILING DATE: 24-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9301764-8
FILING DATE: 24-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky, Martin
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: A1355D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-454-3816
TELEFAX: 610-454-3808

INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: Internal
US-08-476-515A-32

Query Match 76.2%; Score 32; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PPLP 7
DB 4 PPLP 8

RESULT 43
US-08-230-047-17
Sequence 17, Application US/08230047
Patent No. 554109
GENERAL INFORMATION:
APPLICANT: Seafoss III, George H.
APPLICANT: Ivaschenko, Yuri D.
APPLICANT: Jaye, Michael C.
TITLE OF INVENTION: EXPRESSION CLONING OF C-SRC SH3 BINDING
NUMBER OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.0 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/230,047
FILING DATE: 19-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Rosanne
REGISTRATION NUMBER: 32,534
REFERENCE/DOCKET NUMBER: A1465-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3817
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-230-047-17

Query Match 76.2%; Score 32; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PPLP 7
DB 4 PPLP 8

RESULT 44
US-08-602-999A-251
Sequence 251, Application US/08602999A
Patent No. 6184205
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: KAY, Brian K.
APPLICANT: THORN, Judith M.
APPLICANT: OUILLIAM, Lawrence A.
APPLICANT: DER, Channing J.
APPLICANT: FOWLKES, Dana M.
APPLICANT: RIDER, James E.

TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
NUMBER OF INVENTION: ISOLATING AND USING SAME
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,999A
FILING DATE: 16-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864

INFORMATION FOR SEQ ID NO: 251:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-602-999A-251

Query Match 76.2%; Score 32; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PPLP 7
DB 6 PPLP 10

RESULT 45
US-08-602-999A-267
Sequence 267, Application US/08602999A
Patent No. 6184205
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: KAY, Brian K.
APPLICANT: THORN, Judith M.
APPLICANT: OUILLIAM, Lawrence A.
APPLICANT: DER, Channing J.
APPLICANT: FOWLKES, Dana M.
APPLICANT: RIDER, James E.
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,999A
FILING DATE: 16-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 267:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-602-999A-267

Query Match 76.2%; Score 32; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PPLP 7
|||||
DB 6 PPLP 10

RESULT 46
US-08-185-432-12
Sequence 12, Application US/08185432
Patent No. 5750652
GENERAL INFORMATION:
APPLICANT: Atavanis-Tsakonas, Spyridon
APPLICANT: Busseau, Isabelle
APPLICANT: Diederich, Robert J.
APPLICANT: Xu, Tian
APPLICANT: Matsuno, Kenji
TITLE OF INVENTION: DELTA PROTEINS, NUCLEIC ACIDS, AND
TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,432
FILING DATE: 21-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 7326-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-185-432-12

Query Match 76.2%; Score 32; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PPLP 7
|||||
DB 4 PPLP 8

RESULT 47
US-08-630-916A-63
Sequence 63, Application US/08630916A
Patent No. 6011137
GENERAL INFORMATION:
APPLICANT: Pirozzi, Gregorio
APPLICANT: Kay, Brian K.
APPLICANT: Fowlkes, Dana M.
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,916A
FILING DATE: 03-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MISTROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 896-8864/9741
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-630-916A-63

Query Match 76.2%; Score 32; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PPLP 7
|||||

Db 5 PPLP 9

RESULT 48

US-08-630-915A-186
Sequence 186, Application US/08630915A
Patent No. 6309820

GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820H

APPLICANT: KAY, Brian K.

APPLICANT: FOWLES, Dana M.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL

TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND

TITLE OF INVENTION: USING SAME

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630.915A

FILING DATE: 03-APR-1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Mistrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 186:

SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-630-915A-186

Query Match

Best Local Similarity 100.0%; Pred. No. 29;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPLP 7

Db 5 PPLP 9

RESULT 49

US-08-185-432-10

Sequence 10, Application US/08185432

Patent No. 5750652

GENERAL INFORMATION:

APPLICANT: Artavanis-Tsakonas, Spyridon

APPLICANT: Busseau, Isabelle

APPLICANT: Diederich, Robert J.

APPLICANT: Xu, Tian

APPLICANT: Matsuno, Kenji

TITLE OF INVENTION: DETEEX PROTEINS, NUCLEIC ACIDS, AND

TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS

CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/185,432

FILING DATE: 21-JAN-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Mistrock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 7326-006

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-185-432-10

Query Match

Best Local Similarity 100.0%; Pred. No. 31;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPLP 7

Db 5 PPLP 9

RESULT 50

US-08-769-745-31

Sequence 31, Application US/08769745

Patent No. 5955259

GENERAL INFORMATION:

APPLICANT: Holmes, Todd C.

APPLICANT: Levitan, Irwin B.

APPLICANT: Brandeis University

TITLE OF INVENTION: Mechanism for the Regulation of Ion

TITLE OF INVENTION: Channel Activity

FILE REFERENCE: BR096-02

CURRENT APPLICATION NUMBER: US/08/769,745

CURRENT FILING DATE: 1996-12-19

NUMBER OF SEQ ID NOS: 41

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 31

LENGTH: 15

TYPE: PRT

ORGANISM: Homo sapien

US-08-769-745-31

Query Match

Best Local Similarity 100.0%; Pred. No. 31;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PPLP 7

Db 4 PPLP 8

Thu Aug 8 08:12:31 2002

us-09-641-801-1.ra1

Page 20

Search completed: August 8, 2002, 07:48:16
Job time: 134 sec
